arXiv: arXiv:0000.0000

Self-normalized Cramér type moderate deviations for martingales

Xiequan Fan[1,*] **Ion Grama**[2,**] **Quansheng Liu**[2,†] **and Qi-Man Shao**[3,‡]

[1] *Center for Applied Mathematics, Tianjin University, Tianjin 300072, China. e-mail:* [*] fanxiequan@hotmail.com

[2] *Université de Bretagne-Sud, LMBA, UMR CNRS 6205, Campus de Tohannic, 56017 Vannes, France. e-mail:* [**] ion.grama@univ-ubs.fr; [†] quansheng.liu@univ-ubs.fr

[3] *Department of Statistics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong. e-mail:* [‡] qmshao@sta.cuhk.edu.hk

Abstract: Let $(X_i, \mathcal{F}_i)_{i\geq 1}$ be a sequence of martingale differences. Set $S_n = \sum_{i=1}^n X_i$ and $[S]_n = \sum_{i=1}^n X_i^2$. We prove a Cramér type moderate deviation expansion for $\mathbf{P}(S_n/\sqrt{[S]_n} \geq x)$ as $n \to +\infty$. Our results partly extend the earlier work of [Jing, Shao and Wang, 2003] for independent random variables.

Keywords and phrases: Martingales, self-normalized sequences, Cramér's moderate deviations.

1. Introduction

Let $(X_i)_{i\geq 1}$ be a sequence of independent random variables with zero means and finite variances: $\mathbf{E}X_i = 0$ and $0 < \mathbf{E}X_i^2 < \infty$ for all $i \geq 1$. Set

$$S_n = \sum_{i=1}^n X_i, \qquad B_n^2 = \sum_{i=1}^n \mathbf{E}X_i^2, \qquad V_n^2 = \sum_{i=1}^n X_i^2.$$

It is well-known that under the Lindeberg condition the central limit theorem (CLT) holds

$$\sup_{x\in\mathbf{R}} \Big| \mathbf{P}(S_n/B_n \leq x) - \Phi(x) \Big| \to 0 \quad \text{as } n \to \infty,$$

where $\Phi(x)$ denotes the standard normal distribution function. Cramér's moderate deviation expansion stated below gives an estimation of the relative error of $\mathbf{P}(S_n/B_n \geq x)$ to $1-\Phi(x)$. If $(X_i)_{i\geq 1}$ are identically distributed with $\mathbf{E}e^{t_0\sqrt{|X_1|}} < \infty$ for some $t_0 > 0$, then for all $0 \leq x = o(n^{1/6})$ as $n \to \infty$,

$$\frac{\mathbf{P}(S_n/B_n \geq x)}{1-\Phi(x)} = 1 + o(1) \qquad \text{and} \qquad \frac{\mathbf{P}(S_n/B_n \leq -x)}{\Phi(-x)} = 1 + o(1). \tag{1.1}$$

Expansion is available for all $0 \leq x = o(n^{1/2})$ if the moment generating function exists. We refer to Chapter VIII of [Petrov, 1975] for further details on the subject.

However, the limit theorems for self-normalized partial sums of independent random variables have put a new countenance on the classical limit theorems. The study of self-normalized partial sums S_n/V_n originates from Student's t-statistic. Student's t-statistic T_n is defined by

$$T_n = \sqrt{n}\,\overline{X}_n/\widehat{\sigma},$$

where

$$\overline{X}_n = \frac{S_n}{n} \quad \text{and} \quad \widehat{\sigma}^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X}_n)^2}{n-1}.$$

It is known that for all $x \geq 0$,

$$\mathbf{P}\Big(T_n \geq x\Big) = \mathbf{P}\bigg(S_n/V_n \geq x\Big(\frac{n}{n+x^2-1}\Big)^{1/2}\bigg),$$

see [Efron, 1969]. So, if we get an asymptotic bound on the tail probabilities for self-normalized partial sums, then we have an asymptotic bound on the tail probabilities for T_n. [Giné, Götze and Mason, 1997] gave a necessary and sufficient condition for the asymptotic normality. [Bentkus, Bloznelis and Götze, 1996] (see also [Bentkus and Götze, 1996]) obtained the exact Berry-Esseen bound for self-normalized partial sums. [Shao, 1997] established a self-normalized Cramér-Chernoff large deviation without any moment assumptions and [Shao, 1999] proved a self-normalized Cramér moderate deviation theorem under $(2+\rho)$th moments. If $(X_i)_{i\geq 1}$ are independent and identically distributed with $\mathbf{E}|X_1|^{2+\rho} < \infty, \rho \in (0, 1]$, then for all $0 \leq x = o(n^{\rho/(4+2\rho)})$ as $n \to \infty$,

$$\frac{\mathbf{P}(S_n/V_n \geq x)}{1 - \Phi\,(x)} = 1 + o(1). \tag{1.2}$$

For symmetric independent random variables with finite third moments, [Wang and Jing, 1999] derived an exponential nonuniform Berry-Esseen bound, while [Chistyakov and Götze, 2003] further refined Wang and Jing's result and obtained the following Cramér type moderate deviation expansion:

$$\frac{\mathbf{P}(S_n/V_n \geq x)}{1 - \Phi(x)} = 1 + O(1)(1+x)^3 B_n^{-3} \sum_{i=1}^{n} \mathbf{E}|X_i|^3, \tag{1.3}$$

where $O(1)$ is bounded by an absolute constant. The expansion (1.3) was further extended to independent but not necessarily identically distributed random variables by [Jing, Shao and Wang, 2003] under finite $(2+\rho)$th moments, $\rho \in (0, 1]$, showing that

$$\frac{\mathbf{P}(S_n/V_n \geq x)}{1 - \Phi\,(x)} = \exp\Big\{O\big(1\big)(1+x)^{2+\rho}\varepsilon_n^{\rho}\Big\} \tag{1.4}$$

uniformly for $0 \leq x = o(\min\{\varepsilon_n^{-1},\ \kappa_n^{-1}\})$, where

$$\varepsilon_n^{\rho} = \sum_{i=1}^{n} \mathbf{E}|X_i|^{2+\rho}/B_n^{2+\rho} \quad \text{and} \quad \kappa_n^2 = \max_{1\leq i\leq n} \mathbf{E}X_i^2/B_n^2. \tag{1.5}$$

For further self-normalized Cramér type moderate deviation results for independent random variables we refer, for example, to [Hu, Shao and Wang, 2009], [Liu, Shao and Wang, 2013], and [Shao and Zhou, 2016]. We also refer to [de la Peña, Lai and Shao, 2009] and [Shao and Wang, 2013] for recent developments in this area.

The theory for self-normalized sums of independent random variables has been studied in depth. However, we are not aware of any such results for martingales. For some closely related topic, that is, exponential inequalities for self-normalized martingales, we refer to [de la Peña, 1999], [Bercu and Touati, 2008], [Chen, Wang, Xu and Miao, 2014] and [Bercu, Delyon and Rio, 2015]. The main purpose of this paper is to establish self-normalized Cramér type moderate deviation results for martingales. Let $(\delta_n)_{n\geq 1}, (\varepsilon_n)_{n\geq 1}$ and $(\kappa_n)_{n\geq 1}$ be three sequences of nonnegative numbers, such that $\delta_n \to 0$, $\varepsilon_n \to 0$ and $\kappa_n \to 0$ as $n \to \infty$. Let $(X_i, \mathcal{F}_i)_{i\geq 1}$ be a sequence of martingale differences satisfying

$$\Big| \sum_{i=1}^{n} \mathbf{E}[X_i^2|\mathcal{F}_{i-1}] - B_n^2 \Big| \leq \delta_n^2 B_n^2,$$

$$\sum_{i=1}^{n} \mathbf{E}[|X_i|^{2+\rho}|\mathcal{F}_{i-1}] \leq \varepsilon_n^{\rho} B_n^{2+\rho},$$

and

$$\max_{1\leq i\leq n} \mathbf{E}[X_i^2|\mathcal{F}_{i-1}] \leq \kappa_n^2 B_n^2\,,$$

where $\rho \in (0, \frac{3}{2}]$. From Corollary 2.1 we have

$$\mathbf{P}(S_n/V_n \geq x) = (1-\Phi(x))(1+o(1)) \tag{1.6}$$

uniformly for $0 \leq x = o(\,\min\{\varepsilon_n^{-\rho/(3+\rho)},\ \delta_n^{-1},\ \kappa_n^{-1}\})$ as $n \to \infty$. A more general Cramér type expansion is obtained in a larger range in our Theorem 2.1, from which we derive a moderate deviation principle for self-normalized martingales. Moreover, when the condition $\sum_{i=1}^{n} \mathbf{E}[|X_i|^{2+\rho}|\mathcal{F}_{i-1}] \leq \varepsilon_n^{\rho} B_n^{2+\rho}$ is replaced by a slightly stronger condition

$$\mathbf{E}[|X_i|^{2+\rho}|\mathcal{F}_{i-1}] \leq (\varepsilon_n B_n)^{\rho}\mathbf{E}[X_i^2|\mathcal{F}_{i-1}],$$

equality (1.6) holds for a larger range of $0 \leq x = o(\,\min\{\varepsilon_n^{-\rho/(4+2\rho)},\ \delta_n^{-1}\})$ for $\rho \in (0, 1]$, see Corollary 2.4. Clearly, our results recover (1.2) for i.i.d. random variables.

The rest of the paper is organized as follows. Our main results are stated and discussed in Section 2. Section 3 provides the preliminary lemmas that are used in the proofs of the main results. In Section 4, we prove the main results.

Throughout the paper the symbols c and c_α, probably supplied with some indices, denote respectively a generic positive absolute constant and a generic positive constant depending only on α.

2. Main results

Let $(X_i, \mathcal{F}_i)_{i=0,...,n}$ be a sequence of martingale differences defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where $X_0 = 0$ and $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq ... \subseteq \mathcal{F}_n \subseteq \mathcal{F}$ are increasing σ-fields. Set

$$S_0 = 0, \qquad S_k = \sum_{i=1}^{k} X_i, \quad k = 1, ..., n. \tag{2.1}$$

Then $S = (S_k, \mathcal{F}_k)_{k=0,...,n}$ is a martingale. Denote $B_n^2 = \sum_{i=1}^n \mathbf{E}X_i^2$. Let $[S]$ and $\langle S \rangle$ be, respectively, the squared variance and the conditional variance of the martingale S, that is

$$[S]_0 = 0, \qquad [S]_k = \sum_{i=1}^{k} X_i^2, \quad k = 1, ..., n,$$

and

$$\langle S \rangle_0 = 0, \qquad \langle S \rangle_k = \sum_{i=1}^{k} \mathbf{E}[X_i^2|\mathcal{F}_{i-1}], \quad k = 1, ..., n. \tag{2.2}$$

In the sequel, we use the following conditions:

(A1) There exists $\delta_n \in [0, \frac{1}{4}]$ such that

$$\Big| \sum_{i=1}^{n} \mathbf{E}[X_i^2|\mathcal{F}_{i-1}] - B_n^2 \Big| \leq \delta_n^2 B_n^2;$$

(A2) There exist $\rho > 0$ and $\varepsilon_n \in (0, \frac{1}{4}]$ such that

$$\sum_{i=1}^{n} \mathbf{E}[|X_i|^{2+\rho}|\mathcal{F}_{i-1}] \leq \varepsilon_n^\rho B_n^{2+\rho};$$

(A3) There exists $\kappa_n \in (0, \frac{1}{4}]$ such that

$$\mathbf{E}[X_i^2|\mathcal{F}_{i-1}] \leq \kappa_n^2 B_n^2, \quad 1 \leq i \leq n;$$

(A4) There exist $\rho \in (0, 1]$ and $\gamma_n \in (0, \frac{1}{4}]$ such that

$$\mathbf{E}[|X_i|^{2+\rho}|\mathcal{F}_{i-1}] \leq (\gamma_n B_n)^\rho \, \mathbf{E}[X_i^2|\mathcal{F}_{i-1}], \quad 1 \leq i \leq n.$$

When $\rho \in (0, 1]$ and $\gamma_n \leq (16/17)^{1/\rho}/4$, conditions (A1) and (A4) imply condition (A2) with $\varepsilon_n = (17/16)^{1/\rho}\gamma_n$. Thus, we may assume that $\varepsilon_n = O(1)\gamma_n$ as $n \to \infty$. It is also easy to see that condition (A4) implies condition (A3) with $\kappa_n = \gamma_n$, see Lemma 3.5.

In practice, we usually have $\max\{\delta_n, \varepsilon_n, \gamma_n, \kappa_n\} \to 0$ as $n \to \infty$. In the case of sums of i.i.d. random variables, conditions (A1), (A2), (A3), and (A4) are satisfied with $\delta_n = 0$, $\varepsilon_n, \gamma_n, \kappa_n = O(\frac{1}{\sqrt{n}})$.

Our first main result is the following Cramér type moderate deviation for the self-normalized martingale

$$W_n = S_n/\sqrt{[S]_n},$$

under conditions (A1), (A2), and (A3).

Theorem 2.1. *Assume that conditions (A1), (A2), and (A3) are satisfied. Set*

$$\rho_1 = \min\{\rho,\, 1\}.$$

Then for all $0 \le x = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$\frac{\mathbf{P}(W_n \ge x)}{1-\Phi\left(x\right)} = \exp\left\{\theta c_\rho\Big(x^{2+\rho_1}\varepsilon_n^{\rho_1} + x^2\delta_n^2 + (1+x)\big(\varepsilon_n^{\rho/(3+\rho)} + \delta_n\big)\Big)\right\}. \quad (2.3)$$

Moreover, the equality remains valid when $\frac{\mathbf{P}(W_n\ge x)}{1-\Phi(x)}$ *is replaced by* $\frac{\mathbf{P}(W_n\le -x)}{\Phi(-x)}$.

Under condition (A2) the best Berry-Esseen bound for standardized martingales is provided by [Haeusler, 1988]. Assuming $\langle S\rangle_n = B_n^2$ a.s., Haeusler proved that

$$\sup_x \Big|\mathbf{P}(S_n/B_n \le x) - \Phi\left(x\right)\Big| \le C\Big(\sum_{i=1}^n \mathbf{E}|X_i/B_n|^{2+\rho}\Big)^{1/(3+\rho)}.$$

Moreover, it was showed that this bound cannot be improved for martingales with finite $(2+\rho)$th moments. In fact, there exist positive absolute constant c and a sequence of martingale differences satisfying $\mathbf{P}(S_n \le 0) - \Phi\left(0\right) \ge c\left(\sum_{i=1}^n \mathbf{E}|X_i/B_n|^{2+\rho}\right)^{1/(3+\rho)}$ for all large enough n. In particular, under conditions (A2) and $\langle S\rangle_n = B_n^2$ a.s., Haeusler's result implies that

$$\sup_x \Big|\mathbf{P}(S_n/B_n \le x) - \Phi\left(x\right)\Big| \le C\varepsilon_n^{\rho/(3+\rho)}. \quad (2.4)$$

Notice that Theorem 2.1 implies that

$$\sup_x \big|\mathbf{P}(W_n \le x) - \Phi\left(x\right)\big| \le C\big(\varepsilon_n^{\rho/(3+\rho)} + \delta_n\big). \quad (2.5)$$

Under conditions (A2) and $\langle S\rangle_n = B_n^2$ a.s., the Berry-Esseen bound in (2.5) for self-normalized martingales is of the same order as the Berry-Esseen bound in (2.4) for standardized martingales.

From Theorem 2.1, we obtain the following result about the equivalence to the normal tail.

Corollary 2.1. *Assume that conditions (A1), (A2), and (A3) are satisfied with* $\rho \in (0, \frac{3}{2}]$. *Then*

$$\frac{\mathbf{P}(W_n \ge x)}{1-\Phi\left(x\right)} = 1 + o(1) \qquad \textit{and} \qquad \frac{\mathbf{P}(W_n \le -x)}{\Phi\left(-x\right)} = 1 + o(1)$$

uniformly for $0 \le x = o(\,\min\{\varepsilon_n^{-\rho/(3+\rho)},\, \kappa_n^{-1},\, \delta_n^{-1}\})$ *as* $n \to \infty$.

Theorem 2.1 also implies the following moderate deviation principles (MDP) for self-normalized martingales.

Corollary 2.2. *Assume conditions (A1), (A2), and (A3) with* $\max\{\delta_n, \varepsilon_n, \kappa_n\} \to 0$ *as* $n \to \infty$*. Let* a_n *be any sequence of real numbers satisfying* $a_n \to \infty$ *and* $a_n\varepsilon_n \to 0$ *as* $n \to \infty$*. Then for each Borel set* B*,*

$$\begin{aligned} -\inf_{x\in B^o}\frac{x^2}{2} &\leq \liminf_{n\to\infty}\frac{1}{a_n^2}\ln\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big) \\ &\leq \limsup_{n\to\infty}\frac{1}{a_n^2}\ln\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big) \leq -\inf_{x\in\overline{B}}\frac{x^2}{2}, \end{aligned} \tag{2.6}$$

where B^o *and* $\overline{B}$ *denote the interior and the closure of* B*, respectively.*

The last corollary shows that the convergence speed of MDP depends only on ε_n and it has nothing to do with the convergence speeds of κ_n and δ_n.

For i.i.d. random variables, the self-normalized MDP was established by [Shao, 1997]. See also [Jing, Liang and Zhou, 2012] for non-identically distributed random variables.

The other main results concern some improvements of Theorem 2.1 when condition (A3) is replaced by the stronger condition (A4). Theorems 2.2 and 2.3 below give respectively lower and upper bounds, while Theorem 2.4 gives a Cramér type moderate deviation expansion sharper than that in Theorem 2.1.

Theorem 2.2. *Assume that conditions (A1), (A2), and (A4) are satisfied.*

[i] *If* $\rho \in (0,1)$*, then for all* $0 \leq x = o(\gamma_n^{-1})$*,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} \geq \exp\bigg\{-c_\rho\Big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)\left(x^\rho\gamma_n^\rho + \gamma_n^\rho + \delta_n\right)\Big)\bigg\}. \tag{2.7}$$

[ii] *If* $\rho = 1$*, then for all* $0 \leq x = o(\gamma_n^{-1})$*,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} \geq \exp\bigg\{-c\Big(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\left(x\gamma_n + \gamma_n|\ln\gamma_n| + \delta_n\right)\Big)\bigg\}. \tag{2.8}$$

Moreover, the two equalities above remain valid when $\frac{\mathbf{P}(W_n\geq x)}{1-\Phi(x)}$ *is replaced by* $\frac{\mathbf{P}(W_n\leq -x)}{\Phi(-x)}$*.*

For any sequence of positive numbers $(\alpha_n)_{n\geq 1}$ denote

$$\widehat{\alpha}_n(x,\rho) = \frac{\alpha_n^{\rho(2-\rho)/4}}{1+x^{\rho(2+\rho)/4}}. \tag{2.9}$$

Theorem 2.3. *Assume that conditions (A1), (A2), and (A4) are satisfied.*

[i] *If* $\rho \in (0,1)$*, then for all* $0 \leq x = o(\gamma_n^{-1})$*,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} \leq \exp\bigg\{c_\rho\Big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)\Big(x^\rho\gamma_n^\rho + \gamma_n^\rho + \delta_n + \widehat{\varepsilon}_n(x,\rho)\Big)\Big)\bigg\}.$$

[ii] *If $\rho = 1$, then for all $0 \leq x = o(\gamma_n^{-1})$,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} \leq \exp\left\{ c\left(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\Big(x\gamma_n + \gamma_n|\ln\gamma_n| + \delta_n + \widehat{\varepsilon}_n(x,1)\Big)\right)\right\}.$$

Moreover, the two equalities above remain valid when $\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)}$ is replaced by $\frac{\mathbf{P}(W_n \leq -x)}{\Phi(-x)}$.

Combining Theorems 2.2 and 2.3, we obtain the following Cramér type moderate deviation expansion for self-normalized martingales under conditions (A1), (A2), and (A4), which is stronger than the expansion in Theorem 2.1 since the term $\varepsilon_n^{\rho/(3+\rho)}$ therein is improved to a smaller one. In what follows, θ stands for values satisfying $|\theta| \leq 1$.

Theorem 2.4. *Assume that conditions (A1), (A2), and (A4) are satisfied.*

[i] *If $\rho \in (0,1)$, then for all $0 \leq x = o(\gamma_n^{-1})$,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = \exp\left\{ \theta c_\rho\left(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)\Big(x^\rho\gamma_n^\rho + \gamma_n^\rho + \delta_n + \widehat{\varepsilon}_n(x,\rho)\Big)\right)\right\}.$$

[ii] *If $\rho = 1$, then for all $0 \leq x = o(\gamma_n^{-1})$,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = \exp\left\{ \theta c\Big(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\Big(x\gamma_n + \gamma_n|\ln\gamma_n| + \delta_n + \widehat{\varepsilon}_n(x,1)\Big)\Big)\right\}.$$

Moreover, the two equalities above remain valid when $\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)}$ is replaced by $\frac{\mathbf{P}(W_n \leq -x)}{\Phi(-x)}$.

Notice that condition (A4) implies condition (A2) with $\varepsilon_n = \gamma_n$. Therefore, it follows from Theorem 2.4 that:

Corollary 2.3. *Assume that conditions (A1) and (A4) are satisfied.*

[i] *If $\rho \in (0,1)$, then for all $0 \leq x = o(\gamma_n^{-1})$,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = \exp\left\{ \theta c_\rho\left(x^{2+\rho}\gamma_n^\rho + x^2\delta_n^2 + (1+x)\Big(\delta_n + \widehat{\gamma}_n(x,\rho)\Big)\right)\right\}.$$

[ii] *If $\rho = 1$, then for all $0 \leq x = o(\gamma_n^{-1})$,*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = \exp\left\{ \theta c\Big(x^3\gamma_n + x^2\delta_n^2 + (1+x)\Big(\delta_n + \gamma_n|\ln\gamma_n| + \widehat{\gamma}_n(x,1)\Big)\Big)\right\}.$$

Moreover, the two equalities above remain valid when $\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)}$ is replaced by $\frac{\mathbf{P}(W_n \leq -x)}{\Phi(-x)}$.

From Theorem 2.4, we also obtain the following result about the equivalence to the normal tail.

Corollary 2.4. *Assume conditions (A1), (A2), and (A4) with* $\rho \in (0,1]$. *Then*

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = 1+o(1) \qquad \textit{and} \qquad \frac{\mathbf{P}(W_n \leq -x)}{\Phi(-x)} = 1+o(1) \tag{2.10}$$

uniformly for $0 \leq x = o(\min\{\varepsilon_n^{-\rho/(2+\rho)},\ \gamma_n^{-\rho/(1+\rho)},\ \delta_n^{-1}\})$ *as* $n \to \infty$.

In the case of i.i.d. random variables, conditions (A1), (A2), and (A4) are satisfied with $\varepsilon_n, \gamma_n = O(1/\sqrt{n})$ and $\delta_n = 0$. Thus, the range $0 \leq x = o(\min\{\varepsilon_n^{-\rho/(2+\rho)},\ \delta_n^{-1},\ \gamma_n^{-\rho/(1+\rho)}\})$ reduces to $0 \leq x = o(n^{-\rho/(4+2\rho)}), n \to \infty$, which is the best possible result such that (2.10) holds (see [Shao, 1999]). Moreover, from Theorem 2.4, we can get the estimation of the rate of convergence in (2.10); for example, when $\rho = 1$ we have:

Corollary 2.5. *Assume conditions (A1), (A2), and (A4) with* $\rho = 1, \varepsilon_n, \gamma_n, \delta_n = O(1/\sqrt{n})$. *Then, for* $x = x_0 n^{\frac{1}{2}-a}$ *with* $0 < a < \frac{4}{11}$ *and* $x_0 > 0$ *fixed, as* $n \to \infty$,

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = \exp\Big\{O(1)\frac{x^3}{\sqrt{n}}\Big\} \quad \textit{and} \quad \frac{\mathbf{P}(W_n \leq -x)}{\Phi(-x)} = \exp\Big\{O(1)\frac{x^3}{\sqrt{n}}\Big\}. \tag{2.11}$$

In particular, for $x = x_0 n^{\frac{1}{6}-b}$ *with* $0 < b < \frac{1}{33}$ *and* $x_0 > 0$ *fixed, as* $n \to \infty$,

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = 1+O\Big(\frac{x^3}{\sqrt{n}}\Big) \quad \textit{and} \quad \frac{\mathbf{P}(W_n \leq -x)}{\Phi(-x)} = 1+O\Big(\frac{x^3}{\sqrt{n}}\Big). \tag{2.12}$$

Notice that the rate of convergence in (2.11) coincides with that in (1.4) for i.i.d. random variables.

3. Preliminary lemmas

The proofs of Theorems 2.1-2.4 are based on a conjugate multiplicative martingale technique for changing the probability measure which is similar to that of the transformation of [Esscher, 1924]. Our approach is inspired by the earlier work of [Grama and Haeusler, 2000] on Cramér moderate deviations for standardized martingales, and by that of [Shao, 1999], [Jing, Shao and Wang, 2003], who developed techniques for moderate deviations of self-normalized sums of independent random variables. We extend these work by introducing a new choice of the density for the change of measure and refining the approaches in [Shao, 1999] and [Jing, Shao and Wang, 2003] to handle self-normalized martingales. A key point of the proof is a new Berry-Esseen bound for martingales under the changed measure, see Proposition 3.1 below.

Let

$$\xi_i = \frac{X_i}{B_n}, \qquad i = 1, ..., n.$$

Then $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$ is also a sequence of martingale differences. Moreover, for simplicity of notations, set

$$M_k = \sum_{i=1}^{k} \xi_i,$$

$$[M]_k = \sum_{i=1}^k \xi_i^2 \quad \text{and} \quad \langle M \rangle_k = \sum_{i=1}^k \mathbf{E}[\xi_i^2 | \mathcal{F}_{i-1}], \quad k = 1, ..., n.$$

Thus

$$W_n = \frac{S_n}{\sqrt{[S]_n}} = \frac{M_n}{\sqrt{[M]_n}}. \tag{3.1}$$

For any real number λ, consider the *exponential multiplicative martingale* $Z(\lambda) = (Z_k(\lambda), \mathcal{F}_k)_{k=0,...,n}$, where

$$Z_0(\lambda) = 1, \quad Z_k(\lambda) = \prod_{i=1}^k \frac{e^{\zeta_i(\lambda)}}{\mathbf{E}[e^{\zeta_i(\lambda)} | \mathcal{F}_{i-1}]}, \quad k = 1, ..., n$$

with

$$\zeta_i(\lambda) = \lambda \xi_i - \lambda^2 \xi_i^2 / 2.$$

Thus, for each real number λ and each $k = 1, ..., n$, the random variable $Z_k(\lambda)$ is nonnegative and $\mathbf{E} Z_k(\lambda) = 1$. The last observation allows us to introduce the *conjugate probability measure* $\mathbf{P}_\lambda := \mathbf{P}_{\lambda,n}$ on $(\Omega, \mathcal{F})$ defined by

$$d\mathbf{P}_\lambda = Z_n(\lambda) d\mathbf{P}. \tag{3.2}$$

Although $(M_k, \mathcal{F}_k)_{k=0,...,n}$ is a martingale under the measure $\mathbf{P}$, it is no longer a martingale under the conjugate probability measure $\mathbf{P}_\lambda$. To obtain a martingale under $\mathbf{P}_\lambda$ we have to center the random variables $\zeta_i(\lambda)$. Denote by $\mathbf{E}_\lambda$ the expectation with respect to $\mathbf{P}_\lambda$. Because $Z(\lambda)$ is a uniformly integrable martingale under $\mathbf{P}$, we have

$$\mathbf{E}_\lambda[\zeta] = \mathbf{E}[\zeta Z_i(\lambda)] \tag{3.3}$$

and

$$\mathbf{E}_\lambda[\zeta | \mathcal{F}_{i-1}] = \frac{\mathbf{E}[\zeta e^{\zeta_i(\lambda)} | \mathcal{F}_{i-1}]}{\mathbf{E}[e^{\zeta_i(\lambda)} | \mathcal{F}_{i-1}]} \tag{3.4}$$

for any $\mathcal{F}_i$-measurable random variable ζ that is integrable with respect to $\mathcal{F}_i$. Set

$$b_i(\lambda) = \mathbf{E}_\lambda[\zeta_i(\lambda) | \mathcal{F}_{i-1}], \quad i = 1, \ldots, n,$$
$$\eta_i(\lambda) = \zeta_i(\lambda) - b_i(\lambda), \quad i = 1, \ldots, n,$$

and

$$Y_k(\lambda) = \sum_{i=1}^k \eta_i(\lambda), \quad k = 1, ..., n. \tag{3.5}$$

Then $Y(\lambda) = (Y_k(\lambda), \mathcal{F}_k)_{k=0,...,n}$ is the *conjugate martingale*. The following semimartingale decomposition is well-known:

$$\sum_{i=1}^k \zeta_i(\lambda) = B_k(\lambda) + Y_k(\lambda), \quad k = 1, ..., n, \tag{3.6}$$

where $B(\lambda) = (B_k(\lambda), \mathcal{F}_k)_{k=0,...,n}$ is the *drift process* defined as

$$B_k(\lambda) = \sum_{i=1}^{k} b_i(\lambda), \quad k = 1, ..., n.$$

By the relation between $\mathbf{E}$ and $\mathbf{E}_\lambda$ on $\mathcal{F}_i$, we have

$$b_i(\lambda) = \frac{\mathbf{E}[\zeta_i(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]}{\mathbf{E}[e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]}, \quad i = 1, ..., n. \tag{3.7}$$

It is easy to compute the conditional variance of the conjugate martingale $Y(\lambda)$ under the measure $\mathbf{P}_\lambda$, for $k = 0, ..., n$,

$$\begin{aligned} \langle Y(\lambda)\rangle_k &= \sum_{i=1}^{k} \mathbf{E}_\lambda[\eta_i(\lambda)^2|\mathcal{F}_{i-1}] \\ &= \sum_{i=1}^{k} \mathbf{E}_\lambda[(\zeta_i(\lambda) - b_i(\lambda))^2|\mathcal{F}_{i-1}] \\ &= \sum_{i=1}^{k} \left(\frac{\mathbf{E}[\zeta_i^2(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]}{\mathbf{E}[e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]} - \frac{\mathbf{E}[\zeta_i(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]^2}{\mathbf{E}[e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]^2} \right). \end{aligned} \tag{3.8}$$

In the sequel, we give the upper and lower bounds for $B_n(\lambda)$. To this end, we need the following three useful lemmas. The proof is similar to tat in [Shao, 1999] and [Jing, Shao and Wang, 2003]. Set

$$\widetilde{\varepsilon}_\lambda = \lambda^2 \mathbf{E}[\xi_i^2 \mathbf{1}_{\{|\lambda\xi_i|>1\}}|\mathcal{F}_{i-1}] + \lambda^3 \mathbf{E}[\xi_i^3 \mathbf{1}_{\{|\lambda\xi_i|\le 1\}}|\mathcal{F}_{i-1}], \quad \lambda \ge 0.$$

If $\mathbf{E}[|\xi_i|^{2+\rho}] < \infty$ for $\rho \in [0, 1]$, then it is obvious that

$$\widetilde{\varepsilon}_\lambda \le \lambda^{2+\rho}\mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}], \quad \lambda \ge 0.$$

Lemma 3.1. *For all $\lambda > 0$ and $\tau \in [\frac{1}{8}, 2]$, we have*

$$\mathbf{E}[e^{\lambda\xi_i - \tau\lambda^2\xi_i^2}|\mathcal{F}_{i-1}] = 1 + (\frac{1}{2} - \tau)\lambda^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] + O(1)\widetilde{\varepsilon}_\lambda,$$

where $O(1)$ is bounded by an absolute constant.

Lemma 3.2. *For all $\lambda > 0$, we have*

$$\begin{aligned} \mathbf{E}[e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] &= 1 + O(1)\widetilde{\varepsilon}_\lambda, \\ \mathbf{E}[\zeta_i(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] &= \frac{1}{2}\lambda^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] + O(1)\widetilde{\varepsilon}_\lambda, \\ \mathbf{E}[\zeta_i^2(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] &= \lambda^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] + O(1)\widetilde{\varepsilon}_\lambda, \\ \mathbf{E}[|\zeta_i(\lambda)|^3e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] &= O(1)\widetilde{\varepsilon}_\lambda, \\ (\mathbf{E}[\zeta_i(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}])^2 &= O(1)\widetilde{\varepsilon}_\lambda, \end{aligned}$$

where $O(1)$ is bounded by an absolute constant.

Lemma 3.3. *Let $Z_i = \xi_i^2 - \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]$. Then for all $\lambda > 0$,*

$$\begin{aligned}\mathbf{E}[Z_i e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] &= O(1)\frac{1}{\lambda^2}\widetilde{\varepsilon}_\lambda, \\ \mathbf{E}[Z_i^2 e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] &= O(1)\frac{1}{\lambda^4}\widetilde{\varepsilon}_\lambda,\end{aligned}$$

where $O(1)$ is bounded by an absolute constant.

Using Lemma 3.2, we obtain the following upper and lower bounds for $B_n(\lambda)$.

Lemma 3.4. *Assume conditions (A2) and (A3) with $\rho \in (0, 1]$. Then for all $0 \le \lambda = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,*

$$B_n(\lambda) = \frac{1}{2}\lambda^2\langle M\rangle_n + O(1)\lambda^{2+\rho}\varepsilon_n^\rho, \tag{3.9}$$

where $O(1)$ is bounded by an absolute constant.

Proof. According to the definition of $b_i(\lambda)$, we have

$$b_i(\lambda) = \frac{\mathbf{E}[\zeta_i(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]}{\mathbf{E}[e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]}.$$

By Lemma 3.2, it follows that

$$\mathbf{E}[\zeta_i(\lambda)e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] = \frac{1}{2}\lambda^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] + O(1)\widetilde{\varepsilon}_\lambda$$

and

$$\mathbf{E}[e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}] = 1 + O(1)\widetilde{\varepsilon}_\lambda. \tag{3.10}$$

Therefore, conditions (A2) and (A3) imply that for all $0 \le \lambda = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$b_i(\lambda) = \frac{1}{2}\lambda^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] + O(1)\widetilde{\varepsilon}_\lambda$$

and

$$B_n(\lambda) = \frac{1}{2}\lambda^2\langle M\rangle_n + O(1)\lambda^{2+\rho}\varepsilon_n^\rho$$

as desired. □

The following lemma shows that condition (A4) implies condition (A3) with $\kappa_n = \gamma_n$.

Lemma 3.5. *Assume condition (A4). Then $\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] \le \gamma_n^2$.*

Proof. By Jensen's inequality and condition (A4), it holds that

$$\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]^{(2+\rho)/2} \le \mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}] \le \gamma_n^\rho\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}],$$

from which we get $\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] \le \gamma_n^2$. □

Lemma 3.6. *Assume condition (A4). Then for any $t \in [0, \rho)$,*

$$\mathbf{E}[|\xi_i|^{2+t}|\mathcal{F}_{i-1}] \leq \gamma_n^t \, \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]. \tag{3.11}$$

Proof. Let l, p, q be defined by the following equations

$$lp = 2, \quad (2+t-l)q = 2+\rho, \quad p^{-1}+q^{-1}=1, \quad l>0, \text{ and } p,q \geq 1.$$

Solving the last equations, we get

$$l = \frac{2(\rho - t)}{\rho}, \qquad p = \frac{\rho}{\rho - t}, \qquad q = \frac{\rho}{t}.$$

By Hölder's inequality and condition (A4), it is easy to see that

$$\begin{aligned}
\mathbf{E}[|\xi_i|^{2+t}|\mathcal{F}_{i-1}] &= \mathbf{E}[|\xi_i|^l|\xi_i|^{2+t-l}|\mathcal{F}_{i-1}] \\
&\leq (\mathbf{E}[|\xi_i|^{lp}|\mathcal{F}_{i-1}])^{1/p}(\mathbf{E}[|\xi_i|^{(2+t-l)q}|\mathcal{F}_{i-1}])^{1/q} \\
&\leq (\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}])^{1/p}(\mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}])^{1/q} \\
&\leq (\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}])^{1/p}(\gamma_n^\rho \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}])^{1/q} \\
&\leq \gamma_n^{\rho/q}\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] \\
&= \gamma_n^t \, \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}].
\end{aligned}$$

This completes the proof of the lemma. □

Lemma 3.7. *Assume conditions (A1) and (A2). Then for any $t \in [0, \rho)$,*

$$\sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+t}|\mathcal{F}_{i-1}] \leq 2\, \varepsilon_n^t. \tag{3.12}$$

Proof. Recall the notations in the proof of Lemma 3.6. It is easy to see that

$$\sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+t}|\mathcal{F}_{i-1}] \leq \sum_{i=1}^n (\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}])^{1/p}(\mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}])^{1/q}.$$

Using Hölder's inequality and conditions (A1) and (A2), we have

$$\begin{aligned}
\sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+t}|\mathcal{F}_{i-1}] &\leq \Big(\sum_{i=1}^n \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]\Big)^{1/p}\Big(\sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]\Big)^{1/q} \\
&\leq 2\, \varepsilon_n^t,
\end{aligned}$$

which gives the desired inequality. □

We will also need the following two lemmas.

Lemma 3.8. *Assume condition (A1). Then for all $x > 0$,*

$$\mathbf{P}\Big(M_n \geq x\sqrt{[M]_n}, \ [M]_n \geq 16\Big) \leq \frac{2}{3}x^{-2/3}\exp\Big\{-\frac{3}{4}x^2\Big\}.$$

Proof. By inequality (11) of [Delyon, 2009], we have for all $\lambda \in \mathbf{R}$,

$$\mathbf{E}\exp\Big\{\lambda M_n - \frac{\lambda^2}{2}\big(\frac{1}{3}[M]_n + \frac{2}{3}\langle M\rangle_n\big)\Big\} \le 1.$$

Applying the last inequality to the exponential inequality of [de la Peña and Pang, 2009] with $p = q = 2$, we deduce that for all $x > 0$,

$$\mathbf{P}\left(\frac{|M_n|}{\sqrt{\frac{3}{2}(\frac{1}{3}[M]_n + \frac{2}{3}\langle M\rangle_n + \mathbf{E}M_n^2)}} \ge x\right) \le \Big(\frac{2}{3}\Big)^{2/3} x^{-2/3}\exp\Big\{-\frac{1}{2}x^2\Big\}. \quad (3.13)$$

By condition (A1) and the fact $\mathbf{E}\langle M\rangle_n = \mathbf{E}M_n^2 = 1$, it is easy to see that for all $x > 0$,

$$\begin{aligned}
\mathbf{P}\Big(M_n \ge x\sqrt{[M]_n},\ [M]_n \ge 16\Big) &\le \mathbf{P}\Big(M_n \ge x\sqrt{\frac{3}{4}[M]_n + 4},\ [M]_n \ge 16\Big)\\
&\le \mathbf{P}\Big(M_n \ge x\sqrt{\frac{3}{4}[M]_n + \frac{3}{2}\langle M\rangle_n + \frac{9}{4}\mathbf{E}M_n^2},\ [M]_n \ge 16\Big)\\
&\le \mathbf{P}\Big(M_n \ge x\sqrt{\frac{3}{4}[M]_n + \frac{3}{2}\langle M\rangle_n + \frac{9}{4}\mathbf{E}M_n^2}\Big)\\
&= \mathbf{P}\Big(M_n \ge \sqrt{\frac{3}{2}}x\,\sqrt{\frac{1}{2}[M]_n + \langle M\rangle_n + \frac{3}{2}\mathbf{E}M_n^2}\ \Big)\\
&\le \frac{2}{3}x^{-2/3}\exp\Big\{-\frac{3}{4}x^2\Big\}
\end{aligned}$$

as desired. □

Lemma 3.9. *Assume conditions (A1) and (A2). Then*

$$\mathbf{P}\big(|[M]_n - \langle M\rangle_n| \ge 1\big) \le c_\rho\,\big(\varepsilon_n^{(2+\rho)/2} + \varepsilon_n^\rho\,\big).$$

Proof. Notice that $[M]_n - \langle M\rangle_n = \sum_{i=1}^n(\xi_i^2 - \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}])$ is a martingale. For ρ, we distinguish two cases as follows.

When $\rho \in (0, 2]$, by the inequality of [von Bahr and Esseen, 1965], it follows that

$$\begin{aligned}
\mathbf{E}[|[M]_n - \langle M\rangle_n|^{(2+\rho)/2}] &\le \sum_{i=1}^n \mathbf{E}[|\xi_i^2 - \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]|^{(2+\rho)/2}]\\
&\le c_1\sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+\rho}]\\
&\le c_2\,\varepsilon_n^\rho,
\end{aligned}$$

where the last line follows by conditions (A1) and (A2). Hence, by Markov's inequality,

$$\begin{aligned}
\mathbf{P}\big(|[M]_n - \langle M\rangle_n| \ge 1\big) &\le \mathbf{E}[|[M]_n - \langle M\rangle_n|^{(2+\rho)/2}]\\
&\le c_2\,\varepsilon_n^\rho,
\end{aligned}$$

When $\rho > 2$, by Rosenthal's inequality (cf., Theorem 2.12 of [Hall and Heyde, 1980]), Lemma 3.7, and condition (A2), it follows that

$$
\begin{aligned}
&\mathbf{E}[|[M]_n - \langle M\rangle_n|^{(2+\rho)/2}] \\
&\leq c_{\rho,1}\bigg(\mathbf{E}\Big(\sum_{i=1}^n \mathbf{E}[(\xi_i^2 - \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}])^2|\mathcal{F}_{i-1}]\Big)^{(2+\rho)/4} + \sum_{i=1}^n \mathbf{E}|\xi_i^2 - \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]|^{(2+\rho)/2}\bigg) \\
&\leq c_{\rho,2}\bigg(\mathbf{E}\Big(\sum_{i=1}^n \mathbf{E}[\xi_i^4|\mathcal{F}_{i-1}]\Big)^{(2+\rho)/4} + \sum_{i=1}^n \mathbf{E}|\xi_i|^{2+\rho}\bigg) \\
&\leq c_{\rho,3}\left(\varepsilon_n^{(2+\rho)/2} + \varepsilon_n^{\rho}\right). \qquad (3.14)
\end{aligned}
$$

This completes the proof of the lemma. □

Consider the predictable process $\Psi(\lambda) = (\Psi_k(\lambda), \mathcal{F}_k)_{k=0,\ldots,n}$, which is related to the martingale M as follows:

$$\Psi_k(\lambda) = \sum_{i=1}^k \ln \mathbf{E}[e^{\zeta_i(\lambda)}|\mathcal{F}_{i-1}]. \qquad (3.15)$$

By equality (3.10), we easily obtain the following elementary bound for the process $\Psi(\lambda)$.

Lemma 3.10. *Assume conditions (A2) and (A3) with $\rho \in (0, 1]$. Then for all $0 \leq \lambda = o(\min\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,*

$$\Psi_n(\lambda) = O(1)\lambda^{2+\rho}\varepsilon_n^{\rho},$$

where $O(1)$ is bounded by an absolute constant.

In the proofs of Theorems 2.2 and 2.3, we make use of the following assertion, which gives us a rate of convergence in the CLT for the conjugate martingale $Y(\lambda)$ under the probability measure $\mathbf{P}_\lambda$.

Proposition 3.1. *Assume conditions (A1) and (A4).*

[i] *If $\rho \in (0, 1)$, then for all $0 \leq \lambda = o(\gamma_n^{-1})$,*

$$\sup_x \Big|\mathbf{P}_\lambda(Y_n(\lambda)/\lambda \leq x) - \Phi(x)\Big| \leq c_\rho \left(\lambda^\rho \gamma_n^\rho + \gamma_n^\rho + \delta_n\right);$$

[ii] *If $\rho = 1$, then for all $0 \leq \lambda = o(\gamma_n^{-1})$,*

$$\sup_x \Big|\mathbf{P}_\lambda(Y_n(\lambda)/\lambda \leq x) - \Phi(x)\Big| \leq c \left(\lambda\gamma_n + \gamma_n|\ln \gamma_n| + \delta_n\right);$$

with the convention that $Y_n(0)/0 = \sum_{i=1}^n \xi_i$.

Similarly, we have the following Berry-Esseen bound.

Proposition 3.2. *Assume conditions (A1), (A2), and (A3). Then for all* $0 \leq \lambda = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$\sup_x \Big| \mathbf{P}_\lambda(Y_n(\lambda)/\lambda \leq x) - \Phi(x) \Big| \leq c_\rho \left(\lambda^{\rho/2} \gamma_n^{\rho/2} + \varepsilon_n^{\rho/(3+\rho)} + \delta_n \right),$$

with the convention that $Y_n(0)/0 = \sum_{i=1}^n \xi_i$.

The proofs of Propositions 3.1 and 3.2 are much more complicated and we give details in the supplemental article [Fan, Grama, Liu and Shao, 2017].

4. Proof of the main results

We start with the proofs of Theorems 2.2 and 2.3, and conclude with the proof of Theorem 2.1.

4.1. Proof of Theorem 2.2

By (3.1), it is easy to see that

$$\Big\{ S_n \geq x\sqrt{[S]_n} \Big\} = \Big\{ M_n \geq x\sqrt{[M]_n} \Big\} \supseteq \Big\{ M_n \geq \frac{x^2 + \lambda^2 [M]_n}{2\lambda} \Big\} = \Big\{ \sum_{i=1}^n \zeta_i(\lambda) \geq \frac{x^2}{2} \Big\}.$$

For all $0 \leq \lambda = o(\gamma_n^{-1})$, according to (3.2), (3.6) and (3.15), we have the following representation:

$$\begin{aligned} \mathbf{P}\Big(W_n \geq x\Big) &= \mathbf{E}_\lambda\Big[Z_n(\lambda)^{-1} \mathbf{1}_{\{S_n \geq x\sqrt{[S]_n}\}} \Big] \\ &= \mathbf{E}_\lambda\Big[\exp\Big\{ -\sum_{i=1}^n \zeta_i(\lambda) + \Psi_n(\lambda) \Big\} \mathbf{1}_{\{M_n \geq x\sqrt{[M]_n}\}} \Big] \\ &\geq \mathbf{E}_\lambda\Big[\exp\Big\{ -Y_n(\lambda) - B_n(\lambda) + \Psi_n(\lambda) \Big\} \mathbf{1}_{\{\sum_{i=1}^n \zeta_i(\lambda) \geq \frac{x^2}{2}\}} \Big] \\ &= \mathbf{E}_\lambda\Big[\exp\Big\{ -Y_n(\lambda) - B_n(\lambda) + \Psi_n(\lambda) \Big\} \mathbf{1}_{\{Y_n(\lambda) \geq \frac{x^2}{2} - B_n(\lambda)\}} \Big]. \end{aligned}$$

Using Lemmas 3.5, 3.4 and 3.10, we get

$$\begin{aligned} \mathbf{P}\Big(W_n \geq x\Big) \ \geq\ & \mathbf{E}_\lambda\Big[\exp\Big\{ -Y_n(\lambda) - \Big(\frac{1}{2}\lambda^2 \langle M\rangle_n + c_1 \lambda^{2+\rho}\varepsilon_n^\rho\Big) \Big\} \\ & \times \mathbf{1}_{\{Y_n(\lambda) \geq \frac{x^2}{2} - (\frac{1}{2}\lambda^2 \langle M\rangle_n + c_1\lambda^{2+\rho}\varepsilon_n^\rho)\}} \Big]. \end{aligned}$$

Condition (A1) implies that

$$|\langle M\rangle_n - 1| \leq \delta_n^2,$$

and thus

$$\begin{aligned} \mathbf{P}\Big(W_n \geq x\Big) \ \geq\ & \mathbf{E}_\lambda\Big[\exp\Big\{ -Y_n(\lambda) - \Big(\frac{1}{2}\lambda^2 + c_1 \lambda^{2+\rho}\varepsilon_n^\rho\Big)(1+\delta_n^2) \Big\} \\ & \times \mathbf{1}_{\{Y_n(\lambda) \geq \frac{x^2}{2} - (\frac{1}{2}\lambda^2 + c_1\lambda^{2+\rho}\varepsilon_n^\rho)(1+\delta_n^2)\}} \Big]. \end{aligned} \qquad (4.1)$$

Let $\overline{\lambda} = \overline{\lambda}(x)$ be the largest solution of the following equation

$$\Big(\frac{1}{2}\lambda^2 + c_1 \lambda^{2+\rho} \varepsilon_n^\rho\Big)(1+\delta_n^2) = \frac{x^2}{2}.$$

The definition of $\overline{\lambda}$ implies that for all $0 \le x = o(\gamma_n^{-1})$,

$$c_2 \, x \le \overline{\lambda} \le \frac{x}{\sqrt{1+\delta_n^2}} \tag{4.2}$$

and

$$\overline{\lambda} = x + c_3 \theta_0 (x^{1+\rho} \varepsilon_n^\rho + x \delta_n^2), \tag{4.3}$$

where $0 \le \theta_0 \le 1$. From (4.1), we obtain

$$\mathbf{P}\Big(W_n \ge x\Big) \ge \exp\bigg\{ - \Big(\frac{1}{2}\overline{\lambda}^2 + c_1 \overline{\lambda}^{2+\rho} \varepsilon_n^\rho\Big)(1+\delta_n^2)\bigg\} \mathbf{E}_{\overline{\lambda}}\Big[e^{-Y_n(\overline{\lambda})} \mathbf{1}_{\{Y_n(\overline{\lambda}) \ge 0\}}\Big]. \tag{4.4}$$

Setting $F_n(y) = \mathbf{P}_{\overline{\lambda}}(Y_n(\overline{\lambda}) \le y)$, we get

$$\mathbf{P}\Big(W_n \ge x\Big) \ge \exp\bigg\{ - c_4\big(\overline{\lambda}^2 \delta_n^2 + \overline{\lambda}^{2+\rho} \varepsilon_n^\rho\big) - \frac{\overline{\lambda}^2}{2}\bigg\} \int_0^\infty e^{-y} dF_n(y). \tag{4.5}$$

By integration by parts, we have the following bound:

$$\int_0^\infty e^{-y} dF_n(y) \ge \int_0^\infty e^{-y} d\Phi(y/\overline{\lambda}) - 2 \sup_y \Big| F_n(y) - \Phi(y/\overline{\lambda}) \Big|. \tag{4.6}$$

For ρ, we distinguish two cases as follows.

Case 1: If $\rho \in (0,1)$, combining (4.5) and (4.6), by Proposition 3.1, we have for all $0 \le x = o(\gamma_n^{-1})$,

$$\begin{aligned} \mathbf{P}\Big(W_n \ge x\Big) \ \ &\ge \ \ \exp\bigg\{ - c_4\big(\overline{\lambda}^2 \delta_n^2 + \overline{\lambda}^{2+\rho} \varepsilon_n^\rho\big) - \frac{\overline{\lambda}^2}{2}\bigg\} \\ &\quad \times \bigg(\int_0^\infty e^{-\overline{\lambda} y} d\Phi(y) - c_{1,\rho} \Big(\overline{\lambda}^\rho \gamma_n^\rho + \gamma_n^\rho + \delta_n \Big) \bigg). \end{aligned} \tag{4.7}$$

Because

$$e^{-\lambda^2/2} \int_0^\infty e^{-\lambda y} d\Phi(y) = 1 - \Phi\left(\lambda\right) \tag{4.8}$$

and

$$\frac{1}{1+\lambda} e^{-\lambda^2/2} \le \sqrt{2\pi}\Big(1 - \Phi\left(\lambda\right)\Big), \qquad \lambda \ge 0, \tag{4.9}$$

we obtain the following lower bound

$$\begin{aligned} \frac{\mathbf{P}\big(W_n \ge x\big)}{1 - \Phi\left(\overline{\lambda}\right)} \ \ &\ge \ \ \exp\Big\{ -c_4 (\overline{\lambda}^2 \delta_n^2 + \overline{\lambda}^{2+\rho} \varepsilon_n^\rho) \Big\} \Big(1 - c_{2,\rho} \left(1 + \overline{\lambda}\right) (\overline{\lambda}^\rho \gamma_n^\rho + \gamma_n^\rho + \delta_n\,) \Big) \\ &\ge \ \ \exp\Big\{ -c_{3,\rho} \Big(\overline{\lambda}^2 \delta_n^2 + \overline{\lambda}^{2+\rho} \varepsilon_n^\rho + (1+\overline{\lambda})(\overline{\lambda}^\rho \gamma_n^\rho + \gamma_n^\rho + \delta_n\,) \Big) \Big\}, \end{aligned} \tag{4.10}$$

for all $0 \leq \overline{\lambda} \leq \frac{1}{2c_{2,\rho}} \min\{\gamma_n^{-\rho/(1+\rho)}, \delta_n^{-1}\}$.

Next, we consider the case of $\frac{1}{2c_{2,\rho}} \min\{\gamma_n^{-\rho/(1+\rho)}, \delta_n^{-1}\} \leq \overline{\lambda} = o(\gamma_n^{-1})$. Let $K \geq 1$ be an absolute constant, whose exact value is chosen later. It is easy to see that

$$\begin{aligned} \mathbf{E}_{\overline{\lambda}}\Big[e^{-Y_n(\overline{\lambda})}\mathbf{1}_{\{Y_n(\overline{\lambda})\geq 0\}}\Big] &\geq \mathbf{E}_{\overline{\lambda}}\Big[e^{-Y_n(\overline{\lambda})}\mathbf{1}_{\{0\leq Y_n(\overline{\lambda})\leq \overline{\lambda}K\tau\}}\Big] \\ &\geq e^{-\overline{\lambda}K\tau}\mathbf{P}_{\overline{\lambda}}\Big(0 \leq Y_n(\overline{\lambda}) \leq \overline{\lambda}K\tau\Big), \end{aligned} \tag{4.11}$$

where $\tau = \overline{\lambda}^{\rho}\gamma_n^{\rho} + \delta_n$. By Proposition 3.1, we have

$$\begin{aligned} \mathbf{P}_{\overline{\lambda}}\Big(0 \leq Y_n(\overline{\lambda}) \leq \overline{\lambda}K\tau\Big) &\geq \mathbf{P}\Big(0 \leq \mathcal{N}(0,1) \leq K\tau\Big) - c_{4,\rho}\tau \\ &\geq \frac{1}{\sqrt{2\pi}}K\tau e^{-K^2\tau^2/2} - c_{4,\rho}\tau \\ &\geq \Big(\frac{1}{3}K - c_{4,\rho}\Big)\tau. \end{aligned}$$

Letting $K \geq 12c_{4,\rho}$, it follows that

$$\mathbf{P}_{\overline{\lambda}}\Big(0 \leq Y_n(\overline{\lambda}) \leq \overline{\lambda}K\tau\Big) \geq \frac{1}{4}K\tau = \frac{1}{4}K\frac{\overline{\lambda}^{1+\rho}\gamma_n^{\rho} + \overline{\lambda}\delta_n}{\overline{\lambda}}.$$

Choosing

$$K = \max\Big\{12c_{4,\rho}, \frac{4}{\sqrt{\pi}}(2c_{2,\rho})^{1+\rho}\Big\}$$

and taking into account that $\frac{1}{2c_{2,\rho}} \min\{\gamma_n^{-\rho/(1+\rho)}, \delta^{-1}\} \leq \overline{\lambda} = o(\gamma_n^{-1})$, we conclude that

$$\mathbf{P}_{\overline{\lambda}}\Big(0 \leq Y_n(\overline{\lambda}) \leq \overline{\lambda}K\tau\Big) \geq \frac{1}{\sqrt{\pi}\overline{\lambda}}.$$

Because the inequality $\frac{1}{\sqrt{\pi}\lambda}e^{-\lambda^2/2} \geq 1 - \Phi(\lambda)$ is valid for all $\lambda \geq 1$, it follows that for all $\frac{1}{2c_{2,\rho}} \min\{\gamma_n^{-\rho/(1+\rho)}, \delta^{-1}\} \leq \overline{\lambda} = o(\gamma_n^{-1})$,

$$\mathbf{P}_{\overline{\lambda}}\Big(0 \leq Y_n(\overline{\lambda}) \leq K\tau\Big) \geq \Big(1 - \Phi\left(\overline{\lambda}\right)\Big)e^{\overline{\lambda}^2/2}. \tag{4.12}$$

Combining (4.4), (4.11), and (4.12), we obtain

$$\frac{\mathbf{P}(W_n \geq x)}{1 - \Phi\left(\overline{\lambda}\right)} \geq \exp\Big\{ - c_{5,\rho}\left(\overline{\lambda}^2\delta_n^2 + \overline{\lambda}^{2+\rho}\varepsilon_n^{\rho} + (1+\overline{\lambda})(\overline{\lambda}^{\rho}\gamma_n^{\rho} + \gamma_n^{\rho} + \delta_n)\right)\Big\}, \tag{4.13}$$

which is valid for all $\frac{1}{2c_{2,\rho}} \min\{\gamma_n^{-\rho/(1+\rho)}, \delta^{-1}\} \leq \overline{\lambda} = o(\gamma_n^{-1})$.

From (4.10) and (4.13), we get for all $0 \leq \overline{\lambda} = o(\gamma_n^{-1})$,

$$\frac{\mathbf{P}(W_n \geq x)}{1 - \Phi\left(\overline{\lambda}\right)} \geq \exp\Big\{ - c_{6,\rho}\left(\overline{\lambda}^2\delta_n^2 + \overline{\lambda}^{2+\rho}\varepsilon_n^{\rho} + (1+\overline{\lambda})(\overline{\lambda}^{\rho}\gamma_n^{\rho} + \gamma_n^{\rho} + \delta_n)\right)\Big\} \tag{4.14}$$

Next, we substitute x for $\overline{\lambda}$ in the tail of the normal law $1-\Phi(\overline{\lambda})$. By (4.2), (4.3), and (4.9), we get

$$\begin{aligned} 1 \leq \frac{\int_{\overline{\lambda}}^{\infty} \exp\{-t^2/2\}dt}{\int_{x}^{\infty} \exp\{-t^2/2\}dt} &\leq 1+\frac{\int_{\overline{\lambda}}^{x} \exp\{-t^2/2\}dt}{\int_{x}^{\infty} \exp\{-t^2/2\}dt} \\ &\leq 1+c_1 x(x-\overline{\lambda})\exp\Big\{(x^2-\overline{\lambda}^2)/2\Big\} \\ &\leq \exp\Big\{c_2\,(x^2\delta_n^2+x^{2+\rho}\varepsilon_n^{\rho})\Big\} \qquad (4.15) \end{aligned}$$

and hence

$$1-\Phi\left(\overline{\lambda}\right)=(1-\Phi(x))\exp\left\{\theta_1 c\,(x^{2+\rho}\varepsilon_n^{\rho}+x^2\delta_n^2)\right\}. \qquad (4.16)$$

Implementing (4.16) in (4.14) and using (4.2), we obtain for all $0\leq x=o(\gamma_n^{-1})$,

$$\frac{\mathbf{P}\big(W_n\geq x\big)}{1-\Phi\left(x\right)}\geq \exp\bigg\{-c_{7,\rho}\Big(x^{2+\rho}\varepsilon_n^{\rho}+x^2\delta_n^2+(1+x)(x^{\rho}\gamma_n^{\rho}+\gamma_n^{\rho}+\delta_n)\Big)\bigg\},$$

which gives the desired lower bound (2.7).

Case 2: If $\rho=1$, using Proposition 3.1 with $\rho=1$, we have for all $0\leq x=o(\gamma_n^{-1})$,

$$\begin{aligned} \mathbf{P}\Big(W_n\geq x\Big) \ \geq\ & \exp\bigg\{-c_1\big(\overline{\lambda}^2\delta_n^2+\overline{\lambda}^3\varepsilon_n\big)-\frac{\overline{\lambda}^2}{2}\bigg\} \\ & \times\bigg(\int_0^{\infty}e^{-\overline{\lambda}y}d\Phi(y)-c_2\Big(\overline{\lambda}\gamma_n+\gamma_n|\ln\gamma_n|+\delta_n\Big)\bigg), \end{aligned}$$

that is, the term γ_n^{ρ} in inequality (4.7) has been replaced by $\gamma_n|\ln\gamma_n|$. By an argument similar to that of *Case 1*, we obtain the desired lower bound (2.8).

Notice that $(-S_k,\mathcal{F}_k)_{k=0,...,n}$ also satisfies conditions (A1), (A2), and (A4). Thus, the same inequalities hold when $\frac{\mathbf{P}(W_n\geq x)}{1-\Phi(x)}$ is replaced by $\frac{\mathbf{P}(W_n\leq -x)}{\Phi(-x)}$ for all $0\leq x=o(\gamma_n^{-1})$. This completes the proof of Theorem 2.2. □

4.2. Proof of Theorem 2.3

We first prove Theorem 2.3 for all $1\leq x=o(\gamma_n^{-1})$. Observe that

$$\begin{aligned} \mathbf{P}\Big(W_n\geq x\Big) &= \mathbf{P}\Big(W_n\geq x,\,|[M]_n-\langle M\rangle_n|\leq\delta_n+1/(2x)\Big) \\ &\quad+\mathbf{P}\Big(W_n\geq x,\,|[M]_n-\langle M\rangle_n|>\delta_n+1/(2x)\Big). \qquad (4.17) \end{aligned}$$

For the the first term on the right hand side of (4.17), by (3.2) and (3.5) with $\lambda=x$, we have the following representation:

$$\begin{aligned} &\mathbf{P}\Big(W_n\geq x,\,|[M]_n-\langle M\rangle_n|\leq\delta_n+1/(2x)\Big) \\ &=\mathbf{E}_x\Big[Z_n(x)^{-1}\mathbf{1}_{\{M_n\geq x\sqrt{[M]_n},\ |[M]_n-\langle M\rangle_n|\leq\delta_n+1/(2x)\}}\Big] \\ &=\mathbf{E}_x\Big[e^{-Y_n(x)-B_n(x)+\Psi_n(x)}\mathbf{1}_{\big\{xM_n\geq x^2\sqrt{1+[M]_n-1},\ |[M]_n-\langle M\rangle_n|\leq\delta_n+1/(2x)\big\}}\Big]. \end{aligned}$$

By the inequality

$$\sqrt{1+y} \geq 1 + y/2 - y^2/2, \qquad y \geq -1,$$

condition (A1) and Lemma 3.4, we have for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
&\mathbf{P}\Big(W_n \geq x\,,\ |[M]_n - \langle M\rangle_n| \leq \delta_n + 1/(2x)\Big) \\
&\leq \mathbf{E}_x\Big[\exp\Big\{-Y_n(x) - B_n(x) + \Psi_n(x)\Big\} \\
&\qquad \times \mathbf{1}_{\big\{xM_n - \frac{1}{2}x^2[M]_n + \frac{1}{2}x^2([M]_n-1)^2 \geq \frac{1}{2}x^2,\ |[M]_n - \langle M\rangle_n| \leq \delta_n + 1/(2x)\big\}}\Big] \\
&\leq \mathbf{E}_x\Big[\exp\Big\{-Y_n(x) - B_n(x) + \Psi_n(x)\Big\} \\
&\qquad \times \mathbf{1}_{\big\{xM_n - \frac{1}{2}x^2[M]_n + x^2([M]_n - \langle M\rangle_n)^2 + x^2(1-\langle M\rangle_n)^2 \geq \frac{1}{2}x^2,\ |[M]_n - \langle M\rangle_n| \leq \delta_n + 1/(2x)\big\}}\Big] \\
&\leq \mathbf{E}_x\Big[\exp\Big\{-Y_n(x) - B_n(x) + \Psi_n(x)\Big\} \\
&\qquad \times \mathbf{1}_{\big\{Y_n(x) \geq -x^2([M]_n - \langle M\rangle_n)^2 - x^2\delta_n^4 + \frac{1}{2}x^2 - B_n(x),\ |[M]_n - \langle M\rangle_n| \leq \delta_n + 1/(2x)\big\}}\Big] \\
&\leq \mathbf{E}_x\Big[\exp\Big\{-Y_n(x) - B_n(x) + \Psi_n(x)\Big\} \\
&\qquad \times \mathbf{1}_{\big\{Y_n(x) \geq -x^{2+\rho}\varepsilon_n^\rho - x^2\delta_n^4 + \frac{1}{2}x^2 - B_n(x),\ |[M]_n - \langle M\rangle_n| \leq (x\varepsilon_n)^{\rho/2}\big\}}\Big] \\
&\quad + \mathbf{E}_x\Big[\exp\Big\{-Y_n(x) - B_n(x) + \Psi_n(x)\Big\} \\
&\qquad \times \mathbf{1}_{\big\{0 > Y_n(x) \geq -x^2([M]_n - \langle M\rangle_n)^2 - x^2\delta_n^4 + \frac{1}{2}x^2 - B_n(x),\ (x\varepsilon_n)^{\rho/2} < |[M]_n - \langle M\rangle_n| \leq \delta_n + 1/(2x)\big\}}\Big] \\
&\leq \mathbf{E}_x\Big[\exp\Big\{-Y_n(x) - B_n(x) + \Psi_n(x)\Big\} \\
&\qquad \times \mathbf{1}_{\big\{Y_n(x) \geq -c_1(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\big\}}\Big] \\
&\quad + \mathbf{E}_x\Big[\exp\Big\{-Y_n(x) - B_n(x) + \Psi_n(x)\Big\} \\
&\qquad \times \mathbf{1}_{\big\{0 > Y_n(x) \geq -\frac{1}{4} - c_2(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2),\ (x\varepsilon_n)^{\rho/2} < |[M]_n - \langle M\rangle_n| \leq \delta_n + 1/(2x)\big\}}\Big] \\
&:= I_1(x) + I_2(x). \qquad (4.18)
\end{aligned}$$

For $I_1(x)$, by an argument similar to the proof of Theorem 2.2, we get for all $0 \leq x = o(\gamma_n^{-1})$,

$$\frac{I_1(x)}{1-\Phi(x)} \leq \begin{cases} \exp\Big\{c_\rho\Big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)\,(x^\rho\gamma_n^\rho + \gamma_n^\rho + \delta_n)\Big)\Big\} & \text{if } \rho \in (0,1), \\ \exp\Big\{c\Big(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\,(x\gamma_n + \gamma_n|\ln\gamma_n| + \delta_n)\Big)\Big\} & \text{if } \rho = 1. \end{cases} \qquad (4.19)$$

Next, consider the item $I_2(x)$. By condition (A1), Lemmas 3.4 and 3.10, it is obvious that for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned} I_2(x) &\leq \exp\Big\{ -\frac{1}{2}x^2 + c_1\big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2\big)\Big\} \\ &\quad \times \mathbf{E}_x\Big[e^{-Y_n(x)}\mathbf{1}_{\big\{0>Y_n(x)\geq -\frac{1}{4}-c_2(x^{2+\rho}\varepsilon_n^\rho+x^2\delta_n^2),\ (x\varepsilon_n)^{\rho/2}<|[M]_n-\langle M\rangle_n|\big\}}\Big] \\ &\leq \exp\Big\{ -\frac{1}{2}x^2 + c_1\big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2\big)\Big\} \\ &\quad \times \mathbf{E}_x\Big[e^{\frac{1}{4}+c_2(x^{2+\rho}\varepsilon_n^\rho+x^2\delta_n^2)}\mathbf{1}_{\big\{(x\varepsilon_n)^{\rho/2}<|[M]_n-\langle M\rangle_n|\big\}}\Big] \\ &\leq e^{\frac{1}{4}}\exp\Big\{ -\frac{1}{2}x^2 + c_3\big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2\big)\Big\}\mathbf{E}_x\Big[\mathbf{1}_{\big\{(x\varepsilon_n)^{\rho/2}<|[M]_n-\langle M\rangle_n|\big\}}\Big] \end{aligned} \tag{4.20}$$

Denote by $\langle M(x)\rangle_n = \sum_{i=1}^n \mathbf{E}_x[\xi_i^2|\mathcal{F}_{i-1}]$. Notice that $\varepsilon_n = O(1)\gamma_n$. From (3.4), using (3.10), Lemmas 3.3, 3.5 and condition (A2), we obtain for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned} &\Big|\langle M(x)\rangle_n - \langle M\rangle_n\Big| \\ &\leq \sum_{i=1}^n \bigg|\frac{\mathbf{E}[\xi_i^2 e^{x\xi_i - x^2\xi_i^2/2}|\mathcal{F}_{i-1}]}{\mathbf{E}[e^{x\xi_i-x^2\xi_i^2/2}|\mathcal{F}_{i-1}]} - \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]\bigg| + \sum_{i=1}^n\bigg(\frac{\mathbf{E}[\xi_i e^{x\xi_i-x^2\xi_i^2/2}|\mathcal{F}_{i-1}]^2}{\mathbf{E}[e^{x\xi_i-x^2\xi_i^2/2}|\mathcal{F}_{i-1}]^2}\bigg) \\ &\leq c_1\sum_{i=1}^n\Big(\mathbf{E}[x^\rho|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}] + (\mathbf{E}[x\xi_i^2|\mathcal{F}_{i-1}])^2\Big) \\ &\leq c_1\sum_{i=1}^n\Big(\mathbf{E}[x^\rho|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}] + x^2\mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}](\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}])^{(2-\rho)/2}\Big) \\ &\leq c_2\, x^\rho\varepsilon_n^\rho. \end{aligned} \tag{4.21}$$

Thus, for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned} I_2(x) &\leq e^{\frac{1}{4}}\exp\Big\{ -\frac{1}{2}x^2 + c_3\big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2\big)\Big\}\mathbf{E}_x\Big[\mathbf{1}_{\big\{\frac{1}{2}(x\varepsilon_n)^{\rho/2}<|[M]_n-\langle M(x)\rangle_n|\big\}}\Big] \\ &\leq \frac{4e^{\frac{1}{4}}}{(x\varepsilon_n)^{\rho(2+\rho)/4}}\exp\Big\{ -\frac{1}{2}x^2 + c_3\big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2\big)\Big\}\mathbf{E}_x\big[|[M]_n - \langle M(x)\rangle_n|^{(2+\rho)/2}\big]. \end{aligned}$$

It is obvious that

$$[M]_n - \langle M(x)\rangle_n = \sum_{i=1}^n(\xi_i^2 - \mathbf{E}_x[\xi_i^2|\mathcal{F}_{i-1}]).$$

Thus, $([M]_i - \langle M(x)\rangle_i, \mathcal{F}_i)_{i=0,\dots,n}$ is a martingale with respect to the probability measure $\mathbf{P}_x$. By the inequality of [von Bahr and Esseen, 1965], it follows that for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
\mathbf{E}_x[|[M]_n - \langle M(x)\rangle_n|^{(2+\rho)/2}] &\leq c_1 \sum_{i=1}^n \mathbf{E}_x[|\xi_i^2 - \mathbf{E}_x[\xi_i^2|\mathcal{F}_{i-1}]|^{(2+\rho)/2}] \\
&\leq c_2 \sum_{i=1}^n \mathbf{E}_x[|\xi_i|^{2+\rho}] \\
&= c_2 \sum_{i=1}^n \frac{\mathbf{E}[|\xi_i|^{2+\rho} e^{\zeta_i(x)}|\mathcal{F}_{i-1}]}{\mathbf{E}[e^{\zeta_i(x)}|\mathcal{F}_{i-1}]} \\
&\leq c_3 \varepsilon_n^\rho. \qquad (4.22)
\end{aligned}$$

Hence, for all $1 \leq x = o(\gamma_n^{-1})$,

$$I_2(x) \leq C \frac{\varepsilon_n^{\rho(2-\rho)/4}}{x^{\rho(2+\rho)/4}} \exp\Big\{ -\frac{1}{2}x^2 + c_3\big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2\big)\Big\}. \qquad (4.23)$$

Next, we give an estimation for $\mathbf{P}\Big(W_n \geq x,\, |[M]_n - \langle M\rangle_n| > \delta_n + 1/(2x)\Big)$. It is obvious that

$$\begin{aligned}
&\mathbf{P}\Big(W_n \geq x,\, |[M]_n - \langle M\rangle_n| > \delta_n + 1/(2x)\Big) \\
&\leq \mathbf{P}\Big(W_n \geq x,\, |[M]_n - 1| + |1 - \langle M\rangle_n| > \delta_n + 1/(2x)\Big) \\
&\leq \mathbf{P}\Big(W_n \geq x,\, |[M]_n - 1| > \delta_n/2 + 1/(2x)\Big).
\end{aligned}$$

To estimate the tail probability in the last line, we follow the argument of [Shao and Zhou, 2016]. We have the following decomposition:

$$\begin{aligned}
&\mathbf{P}\Big(W_n \geq x,\; |[M]_n - 1| > \delta_n/2 + 1/(2x)\Big) \\
&\leq \mathbf{P}\Big(M_n/\sqrt{[M]_n} \geq x,\; 1 + \delta_n/2 + 1/(2x) < [M]_n \leq 16\Big) \\
&\quad + \mathbf{P}\Big(M_n/\sqrt{[M]_n} \geq x,\; [M]_n < 1 - \delta_n/2 - 1/(2x)\Big) \\
&\quad + \mathbf{P}\Big(M_n/\sqrt{[M]_n} \geq x,\; [M]_n > 16\Big) \\
&:= \sum_{v=1}^3 \mathbf{P}\Big((M_n,\, \sqrt{[M]_n}) \in \mathcal{E}_v\Big), \qquad (4.24)
\end{aligned}$$

where $\mathcal{E}_v \subset \mathbf{R} \times \mathbf{R}^+, 1 \leq v \leq 3$, are given by

$$\begin{aligned}
\mathcal{E}_1 &= \Big\{(u,v) \in \mathbf{R} \times \mathbf{R}^+ : u/v \geq x,\; \sqrt{1 + \delta_n/2 + 1/(2x)} < v \leq 4\Big\}, \\
\mathcal{E}_2 &= \Big\{(u,v) \in \mathbf{R} \times \mathbf{R}^+ : u/v \geq x,\; v < \sqrt{1 - \delta_n/2 - 1/(2x)}\;\Big\}, \\
\mathcal{E}_3 &= \Big\{(u,v) \in \mathbf{R} \times \mathbf{R}^+ : u/v \geq x,\; v > 4\Big\}.
\end{aligned}$$

To estimate the probability $\mathbf{P}((M_n, \sqrt{[M]_n}) \in \mathcal{E}_1)$, we introduce the following new conjugate probability measure $\widetilde{\mathbf{P}}_x$ defined by

$$d\widetilde{\mathbf{P}}_x = \widetilde{Z}_n(x) d\mathbf{P},$$

where

$$\widetilde{Z}_n(x) = \prod_{i=1}^{k} \frac{e^{\widetilde{\zeta}_i(x)}}{\mathbf{E}[e^{\widetilde{\zeta}_i(x)}|\mathcal{F}_{i-1}]} \quad \text{ and } \quad \widetilde{\zeta}_i(x) = x\xi_i - x^2\xi_i^2/8.$$

Denote by $\widetilde{\mathbf{E}}_x$ the expectation with respect to $\widetilde{\mathbf{P}}_x$ and $\langle \widetilde{M}(x)\rangle_n = \sum_{i=1}^n \widetilde{\mathbf{E}}_x[\xi_i^2|\mathcal{F}_{i-1}]$. By an argument similar to (4.21), it follows that for all $1 \leq x = o(\gamma_n^{-1})$,

$$\langle \widetilde{M}(x)\rangle_n = \langle M\rangle_n + O(1)x^\rho \varepsilon_n^\rho.$$

By Markov's inequality, we deduce that

$$\begin{aligned}
&\mathbf{P}\Big((M_n, \sqrt{[M]_n}) \in \mathcal{E}_1\Big) \\
&\leq (\delta_n/2 + 1/(2x))^{-2} e^{-\inf_{(u,v)\in\mathcal{E}_1}(xu-(vx)^2/8)}\mathbf{E}[([M]_n - 1)^2 e^{xM_n - [M]_n x^2/8}] \\
&\leq 16x^2 e^{-\inf_{(u,v)\in\mathcal{E}_1}(xu-(vx)^2/8)}\mathbf{E}[([M]_n - \langle \widetilde{M}(x)\rangle_n)^2 e^{xM_n - [M]_n x^2/8}] \\
&\quad + 16x^2 e^{-\inf_{(u,v)\in\mathcal{E}_1}(xu-(vx)^2/8)}\mathbf{E}[(\langle \widetilde{M}(x)\rangle_n - \langle M\rangle_n)^2 e^{xM_n - [M]_n x^2/8}] \\
&\quad + 16\delta_n^{-2} e^{-\inf_{(u,v)\in\mathcal{E}_1}(xu-(vx)^2/8)}\mathbf{E}[(\langle M\rangle_n - 1)^2 e^{xM_n - [M]_n x^2/8}] \\
&\leq 16x^2 e^{-\inf_{(u,v)\in\mathcal{E}_1}(xu-(vx)^2/8)}\mathbf{E}[([M]_n - \langle \widetilde{M}(x)\rangle_n)^2 e^{xM_n - [M]_n x^2/8}] \\
&\quad + Cx^{2+2\rho}\varepsilon_n^{2\rho} e^{-\inf_{(u,v)\in\mathcal{E}_1}(xu-(vx)^2/8)}\mathbf{E}[e^{xM_n - [M]_n x^2/8}] \\
&\quad + 16\delta_n^2 e^{-\inf_{(u,v)\in\mathcal{E}_1}(xu-(vx)^2/8)}\mathbf{E}[e^{xM_n - [M]_n x^2/8}], \qquad (4.25)
\end{aligned}$$

where it is easy to verify that

$$\inf_{(u,v)\in\mathcal{E}_1}\Big(xu - \frac{1}{8}(vx)^2\Big) \geq \frac{7}{8}x^2 + \frac{1}{4}x - c\,x^2\delta_n^2. \qquad (4.26)$$

By Lemma 3.1, conditions (A1) and (A2), it follows that

$$\begin{aligned}
\prod_{i=1}^n \mathbf{E}[e^{\widetilde{\zeta}_i(x)}|\mathcal{F}_{i-1}] &\leq \prod_{i=1}^n \Big(1 + \frac{3}{8}x^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] + O(1)x^{2+\rho}\mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]\Big) \\
&\leq \prod_{i=1}^n \exp\Big\{\frac{3}{8}x^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] + O(1)x^{2+\rho}\mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]\Big\} \\
&= \exp\Big\{\frac{3}{8}x^2\langle M\rangle_n + O(1)x^{2+\rho}\sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]\Big\} \\
&\leq \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\}.
\end{aligned}$$

Therefore, for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
&\mathbf{E}\Big[([M]_n - \langle \widetilde{M}(x)\rangle_n)^2 e^{xM_n - [M]_n x^2/8}\Big] \\
&= \mathbf{E}\Big[\Big(\Pi_{i=1}^n \mathbf{E}[e^{\widetilde{\zeta}_i(x)}|\mathcal{F}_{i-1}]\Big)\Big([M]_n - \langle \widetilde{M}(x)\rangle_n\Big)^2 \widetilde{Z}_n(x)\Big] \\
&\leq \mathbf{E}\Big[\Big([M]_n - \langle \widetilde{M}(x)\rangle_n\Big)^2 \widetilde{Z}_n(x)\Big] \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\} \\
&= \widetilde{\mathbf{E}}_x\Big[\Big([M]_n - \langle \widetilde{M}(x)\rangle_n\Big)^2\Big] \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\} \\
&= \sum_{i=1}^n \widetilde{\mathbf{E}}_x\Big[(\xi_i^2 - \widetilde{\mathbf{E}}_x[\xi_i^2|\mathcal{F}_{i-1}])^2\Big] \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\},
\end{aligned}$$

where the last line follows because $([M]_i - \langle \widetilde{M}(x)\rangle_i, \mathcal{F}_i)_{i=0,\dots,n}$ is a martingale with respect to the probability measure $\widetilde{\mathbf{P}}_x$. Therefore, by Lemma 3.1, conditions (A1) and (A2) again, we have for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
&\mathbf{E}\Big[([M]_n - \langle \widetilde{M}(x)\rangle_n)^2 e^{xM_n - [M]_n x^2/8}\Big] \\
&\leq \sum_{i=1}^n \widetilde{\mathbf{E}}_x\big[\widetilde{\mathbf{E}}_x\big[\xi_i^4|\mathcal{F}_{i-1}\big]\big] \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\} \\
&= \sum_{i=1}^n \widetilde{\mathbf{E}}_x\Big[\mathbf{E}[\xi_i^4 e^{\widetilde{\zeta}_i(x)}|\mathcal{F}_{i-1}]\Big/\mathbf{E}[e^{\widetilde{\zeta}_i(x)}|\mathcal{F}_{i-1}]\Big] \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\} \\
&\leq C_0 \sum_{i=1}^n \widetilde{\mathbf{E}}_x\Big[\frac{1}{x^{2-\rho}} \sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]\Big] \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\} \\
&\leq C_1 \varepsilon_n^\rho \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\}.
\end{aligned}$$

Lemma 3.1 implies that for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
&\mathbf{E}\Big[\exp\Big\{xM_n - \frac{1}{8}x^2[M]_n - \frac{3}{8}x^2\langle M\rangle_n - O(1)x^{2+\rho}\sum_{i=1}^n \mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]\Big\}\Big] \\
&\leq \mathbf{E}\Big[\exp\Big\{xM_{n-1} - \frac{1}{8}x^2[M]_{n-1} - \frac{3}{8}x^2\langle M\rangle_{n-1} - O(1)x^{2+\rho}\sum_{i=1}^{n-1} \mathbf{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]\Big\}\Big] \\
&\leq 1.
\end{aligned}$$

By conditions (A1), (A2) and the last inequality, we obtain for all $1 \leq x = o(\gamma_n^{-1})$,

$$\mathbf{E}[e^{xM_n - [M]_n x^2/8}] \leq \exp\Big\{\frac{3}{8}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\}.$$

Thus, from (4.25), we deduce that for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
&\mathbf{P}\Big((M_n, \sqrt{[M]_n}) \in \mathcal{E}_1\Big) \\
&\leq C_2(\varepsilon_n^\rho + x^{2+2\rho}\varepsilon_n^{2\rho} + \delta_n^2) \exp\Big\{ -\frac{1}{2}x^2 - \frac{1}{4}x + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\} \\
&\leq C_3(\varepsilon_n^\rho + \delta_n^2) \exp\Big\{ -\frac{1}{2}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\}. \qquad (4.27)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\mathbf{P}\Big((M_n, \sqrt{[M]_n}) \in \mathcal{E}_2\Big) \\
&\leq (\delta_n/2 + 1/(2x))^{-2} e^{-\inf_{(u,v)\in\mathcal{E}_2}(xu - 2(vx)^2)} \mathbf{E}[([M]_n - 1)^2 e^{xM_n - 2[M]_n x^2}] \\
&\leq C_4(\varepsilon_n^\rho + \delta_n^2) \exp\Big\{ -\frac{1}{2}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\}. \qquad (4.28)
\end{aligned}$$

For the last term $\mathbf{P}((M_n, \sqrt{[M]_n}) \in \mathcal{E}_3)$, we obtain the following estimation

$$\begin{aligned}
\mathbf{P}\Big((M_n, \sqrt{[M]_n}) \in \mathcal{E}_3\Big) &= \mathbf{P}\Big(M_n \geq x\sqrt{[M]_n},\ [M]_n > 16\Big) \\
&\leq \frac{2}{3}x^{-2/3} \exp\Big\{ -\frac{3}{4}x^2\Big\}, \qquad (4.29)
\end{aligned}$$

where the last line follows by Lemma 3.8. Moreover, by Lemma 3.9, it holds that for $\rho \in (0, 1]$,

$$\begin{aligned}
\mathbf{P}\Big((M_n, \sqrt{[M]_n}) \in \mathcal{E}_3\Big) &\leq \mathbf{P}\Big(|[M]_n - \langle M\rangle_n| \geq 1\Big) \\
&\leq c\,\varepsilon_n^\rho.
\end{aligned}$$

By the last inequality and (4.29), we get for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
\mathbf{P}\Big((M_n, \sqrt{[M]_n}) \in \mathcal{E}_3\Big) &\leq \min\Big\{c\,\varepsilon_n^\rho,\ \frac{2}{3}x^{-2/3}e^{-3x^2/4}\Big\} \\
&\leq C\,\frac{\varepsilon_n^{\rho(2-\rho)/4}}{x^{\rho(2+\rho)/4}} \exp\Big\{ -\frac{1}{2}x^2\Big\}. \qquad (4.30)
\end{aligned}$$

Thus, combining the inequalities (4.24), (4.27), (4.28) and (4.30) together, we deduce that for all $1 \leq x = o(\gamma_n^{-1})$,

$$\begin{aligned}
&\mathbf{P}\Big(W_n \geq x,\ |[M]_n - \langle M\rangle_n| > \delta_n + 1/(2x)\Big) \\
&\leq C\Big(\frac{\varepsilon_n^{\rho(2-\rho)/4}}{x^{\rho(2+\rho)/4}} + \delta_n^2\Big) \exp\Big\{ -\frac{1}{2}x^2 + O(1)(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2)\Big\}. \qquad (4.31)
\end{aligned}$$

Combining (4.18), (4.19), (4.23), and (4.31), we obtain for all $1 \leq x = o(\gamma_n^{-1})$,

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} \leq \left(1 + C(1+x)\Big(\frac{\varepsilon_n^{\rho(2-\rho)/4}}{x^{\rho(2+\rho)/4}} + \delta_n^2\Big)\right)$$
$$\times \begin{cases} \exp\left\{c_\rho\Big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)(x^\rho\gamma_n^\rho + \gamma_n^\rho + \delta_n)\Big)\right\} & \text{if } \rho \in (0,1) \\ \exp\left\{c\Big(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\left(x\gamma_n + \gamma_n|\ln\gamma_n| + \delta_n\right)\Big)\right\} & \text{if } \rho = 1 \end{cases}$$
$$\leq \begin{cases} \exp\left\{C_\rho\Big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)\big(x^\rho\gamma_n^\rho + \gamma_n^\rho + \delta_n + \frac{\varepsilon_n^{\rho(2-\rho)/4}}{x^{\rho(2+\rho)/4}}\big)\Big)\right\} & \text{if } \rho \in (0,1) \\ \exp\left\{C\Big(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\big(x\gamma_n + \gamma_n|\ln\gamma_n| + \delta_n + \frac{\varepsilon_n^{\rho(2-\rho)/4}}{x^{\rho(2+\rho)/4}}\big)\Big)\right\} & \text{if } \rho = 1, \end{cases}$$

which gives the desired inequalities.

For the case of $0 \leq x < 1$, the proof of Theorem 2.3 is similar to the case of $x = 1$. Notice that $(-S_k, \mathcal{F}_k)_{k=0,\dots,n}$ also satisfies conditions (A1), (A2), and (A4). Thus, the same inequalities hold when $\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)}$ is replaced by $\frac{\mathbf{P}(W_n \leq -x)}{\Phi(-x)}$ for all $0 \leq x = o(\gamma_n^{-1})$. This completes the proof of Theorem 2.3. □

4.3. Proof of Theorem 2.1

Using Proposition 3.2, by an argument similar to the proof of Theorem 2.4, we obtain the following result. If $\rho \in (0,1)$, then for all $0 \leq x = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)}$$
$$= \exp\left\{\theta c_\rho\left(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)\left(x^{\rho/2}\varepsilon_n^{\rho/2} + \varepsilon_n^{\rho/(3+\rho)} + \delta_n + \frac{\varepsilon_n^{\rho(2-\rho)/4}}{1+x^{\rho(2+\rho)/4}}\right)\right)\right\}.$$

Notice that the following three inequalities hold:

$$\begin{array}{rcll} x^{1+\rho/2}\varepsilon_n^{\rho/2} & \leq & x^{2+\rho}\varepsilon_n^\rho, & x \geq \varepsilon_n^{-\rho/(2+\rho)}, \\ x^{\rho/2}\varepsilon_n^{\rho/2} & \leq & \varepsilon_n^{\rho/(3+\rho)}, & 0 \leq x \leq \varepsilon_n^{-\rho/(2+\rho)}, \\ \varepsilon_n^{\rho(2-\rho)/4} & \leq & \varepsilon_n^{\rho/(3+\rho)}, & \rho \in (0,1]. \end{array}$$

Therefore, for $\rho \in (0,1)$ and all $0 \leq x = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} = \exp\left\{\theta c_\rho\Big(x^{2+\rho}\varepsilon_n^\rho + x^2\delta_n^2 + (1+x)\big(\varepsilon_n^{\rho/(3+\rho)} + \delta_n\big)\Big)\right\},$$

which gives the desired equality for $\rho \in (0,1)$.

Assume that condition (A2) holds for $\rho \geq 1$. When $\rho \in [1, 2]$, by Markov's inequality and (4.22), we have for all $x \geq 1$,

$$\begin{aligned}\mathbf{E}_x\Big[\mathbf{1}_{\{(x\varepsilon_n)^{1/2}<|[M]_n-\langle M\rangle_n|\}}\Big] &\leq \frac{1}{(x\varepsilon_n)^{(2+\rho)/4}}\mathbf{E}_x\big[|[M]_n-\langle M(x)\rangle_n|^{(2+\rho)/2}\big] \\ &\leq \frac{1}{x^{(2+\rho)/4}}\varepsilon_n^{(3\rho-2)/4} \\ &\leq \varepsilon_n^{(3\rho-2)/4}. \qquad (4.32)\end{aligned}$$

When $\rho > 2$, Lemma 3.7 implies that condition (A2) also holds for $\rho = 2$, with the term ε_n in condition (A2) replaced by $2\varepsilon_n$. Then (4.32) with $\rho = 2$ shows that

$$\mathbf{E}_x\Big[\mathbf{1}_{\{(x\varepsilon_n)^{1/2}<|[M]_n-\langle M\rangle_n|\}}\Big] \leq 2\varepsilon_n.$$

Thus, for all $\rho \geq 1$, it holds that

$$\mathbf{E}_x\Big[\mathbf{1}_{\{(x\varepsilon_n)^{1/2}<|[M]_n-\langle M\rangle_n|\}}\Big] \leq \max\Big\{\varepsilon_n^{(3\rho-2)/4},\ 2\varepsilon_n\Big\} \leq 2\varepsilon_n^{\rho/(3+\rho)}.$$

Notice that Lemma 3.7 also implies that condition (A2) holds for $\rho = 1$. Therefore, by (4.20), (4.23) can be improved to

$$\begin{aligned}I_2(x) &\leq e^{\frac{1}{4}}\exp\Big\{-\frac{1}{2}x^2 + c_3\big(x^{2+\rho}\varepsilon_n^{\rho} + x^2\delta_n^2\big)\Big\}\mathbf{E}_x\Big[\mathbf{1}_{\{(x\varepsilon_n)^{1/2}<|[M]_n-\langle M\rangle_n|\}}\Big] \\ &\leq C\,\varepsilon_n^{\rho/(3+\rho)}\exp\Big\{-\frac{1}{2}x^2 + c_3\big(x^3\varepsilon_n + x^2\delta_n^2\big)\Big\}.\end{aligned}$$

Notice also that for $\rho \geq 1$,

$$\begin{aligned}\mathbf{P}\Big((M_n, \sqrt{[M]_n}) \in \mathcal{E}_3\Big) &\leq \min\Big\{c\,\varepsilon_n^{\rho},\ \frac{2}{3}x^{-2/3}e^{-3x^2/4}\Big\} \\ &\leq C\,\varepsilon_n^{\rho/(3+\rho)}\exp\Big\{-\frac{1}{2}x^2\Big\}.\end{aligned}$$

By an argument similar to the proof for case $\rho \in (0, 1)$ but with the term $(x\varepsilon_n)^{\rho/2}$ in (4.18) replaced by $(x\varepsilon_n)^{1/2}$, we have for all $0 \leq x = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$\begin{aligned}\frac{\mathbf{P}(W_n \geq x)}{1-\Phi(x)} &= \exp\Big\{\theta c_1\Big(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\big(x^{\rho/2}\varepsilon_n^{\rho/2} + \varepsilon_n^{\rho/(3+\rho)} + \delta_n\big)\Big)\Big\} \\ &= \exp\Big\{\theta c_2\Big(x^3\varepsilon_n + x^2\delta_n^2 + (1+x)\big(\varepsilon_n^{\rho/(3+\rho)} + \delta_n\big)\Big)\Big\},\end{aligned}$$

which gives the desired equality for $\rho \geq 1$.

4.4. Proof of Corollary 2.2

To prove Corollary 2.2, we need the following two sides bound on the tail probabilities of the standard normal random variable:

$$\frac{1}{\sqrt{2\pi}(1+x)}e^{-x^2/2} \leq 1-\Phi(x) \leq \frac{1}{\sqrt{\pi}(1+x)}e^{-x^2/2}, \qquad x \geq 0. \qquad (4.1)$$

First, we prove that

$$\limsup_{n\to\infty}\frac{1}{a_n^2}\ln\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big)\leq-\inf_{x\in\overline{B}}\frac{x^2}{2}. \tag{4.2}$$

For any given Borel set $B\subset\mathbf{R}$, let $x_0=\inf_{x\in B}|x|$. Then, it is obvious that $x_0\geq\inf_{x\in\overline{B}}|x|$. Therefore, by Theorem 2.1,

$$\begin{aligned}\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big) &\leq \mathbf{P}\Big(\big|W_n\big|\geq a_nx_0\Big)\\ &\leq 2\Big(1-\Phi\left(a_nx_0\right)\Big)\\ &\quad\times\exp\bigg\{c_\rho\bigg(\left(a_nx_0\right)^{2+\rho}\varepsilon_n^\rho+\left(a_nx_0\right)^2\delta_n^2+\left(a_nx_0\right)\left(\varepsilon_n^{\rho/(3+\rho)}+\delta_n\right)\bigg)\bigg\}.\end{aligned}$$

Using (4.1), we deduce that

$$\limsup_{n\to\infty}\frac{1}{a_n^2}\ln\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big)\ \leq\ -\frac{x_0^2}{2}\ \leq\ -\inf_{x\in\overline{B}}\frac{x^2}{2},$$

which gives (4.2).

Next, we prove that

$$\liminf_{n\to\infty}\frac{1}{a_n^2}\ln\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big)\geq-\inf_{x\in B^o}\frac{x^2}{2}. \tag{4.3}$$

We may assume that $B^o\neq\emptyset$. For any $\varepsilon_1>0$, there exists an $x_0\in B^o$, such that

$$0<\frac{x_0^2}{2}\leq\inf_{x\in B^o}\frac{x^2}{2}+\varepsilon_1. \tag{4.4}$$

For $x_0\in B^o$, there exists small $\varepsilon_2\in(0,x_0)$, such that $(x_0-\varepsilon_2,x_0+\varepsilon_2]\subset B$. Then it is obvious that $x_0\geq\inf_{x\in\overline{B}}x$. Without loss of generality, we may assume that $x_0>0$. By Theorem 2.1, we deduce that

$$\begin{aligned}\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big) &\geq \mathbf{P}\Big(W_n\in(a_n(x_0-\varepsilon_2),a_n(x_0+\varepsilon_2)]\Big)\\ &\geq \mathbf{P}\Big(W_n>a_n(x_0-\varepsilon_2)\Big)-\mathbf{P}\Big(W_n>a_n(x_0+\varepsilon_2)\Big).\end{aligned}$$

Using Theorem 2.1 and (4.1), it follows that

$$\liminf_{n\to\infty}\frac{1}{a_n^2}\ln\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big)\geq-\frac{1}{2}(x_0-\varepsilon_2)^2.$$

Letting $\varepsilon_2\to0$, we get

$$\liminf_{n\to\infty}\frac{1}{a_n^2}\ln\mathbf{P}\Big(\frac{W_n}{a_n}\in B\Big)\ \geq\ -\frac{x_0^2}{2}\ \geq\ -\inf_{x\in B^o}\frac{x^2}{2}-\varepsilon_1.$$

Because ε_1 can be arbitrarily small, we obtain (4.3). This completes the proof of Corollary 2.2. ☐

Acknowledgements

Fan and Liu have been partially supported by the National Natural Science Foundation of China (Grant nos. 11601375, 11626250, 11571052, 11401590 and 11731012), and by Hunan Natural Science Foundation (China, grant no. 2017JJ2271). Grama and Liu have benefitted from the support of the French government "Investissements d'Avenir" program ANR-11-LABX-0020-01. Shao has been partially supported by Hong Kong RGC GRF 14302515.

Supplementary Material

Supplement to "Self-normalized Cramér type moderate deviations for martingales"
(doi: COMPLETED BY THE TYPESETTER; .pdf). The supplement gives the detailed proofs of Propositions 3.1 and 3.2.

References

Bercu, B., Delyon, B., Rio, E. (2015). *Concentration inequalities for sums and martingales.* Springer, New York.

Bercu, B., Touati, A. (2008). Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.*, **18**(5): 1848–1869.

Bentkus, V., Bloznelis, M., Götze, F. (1996). A Berry-Esséen bound for Student's statistic in the non-i.i.d. case. *J. Theoret. Probab.*, **9**(3), 765–796.

Bentkus, V., Götze, F. (1996). The Berry-Esseen bound for Student's statistic. *Ann. Probab.*, **24**(1), 491–501.

Chen, S., Wang, Z., Xu, W., Miao, Y. (2014). Exponential inequalities for self-normalized martingales. *J. Inequal. Appl.*, 2014: 289.

Chistyakov, G.P., Götze, F. (2003). Moderate deviations for Student's statistic. *Theory Probab. Appl.*, **47**(3): 415–428.

de la Peña, V.H. (1999). A general class of exponential inequalities for martingales and ratios. *Ann. Probab.*, **27**(1): 537–564.

de la Peña, V.H., Lai, T.L., Shao, Q.M. (2009) *Self-normalized Processes: Theory and Statistical Applications.* Springer Series in Probability and its Applications. Springer-Verlag. New York.

de la Peña, V. H., Pang, G. (2009). Exponential inequalities for self-normalized processes with applications. *Electron. Commun. Probab.*, **14**: 372–382.

Delyon, B. (2009). Exponential inequalities for sums of weakly dependent variables. *Electron. J. Probab.*, **14**(28): 752–779.

Efron, B. (1969). Student's t-test under symmetry conditions, *J. Amer. Statist. Assoc.*, **64**(328): 1278–1302.

Esscher, F. (1924). On a method of determining correlation from the ranks of the variates. *Scand. Actuar. J.*, **1**: 201–219.

Fan, X., Grama, I., Liu, Q., Shao, Q.M. (2017). Supplement to "Self-normalized Cramér type moderate deviations for martingales".

Giné, E., Götze, F., Mason, D.M. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab.*, **25**(3): 1514–1531.

Grama, I., Haeusler, E. (2000). Large deviations for martingales via Cramér's method. *Stochastic Process. Appl.*, **85**(2), 279–293.

Haeusler, E. (1988). On the rate of convergence in the central limit theorem for martingales with discrete and continuous time. *Ann. Probab.*, **16**(1), 275–299.

Hall, P., Heyde, C.C. (1980). *Martingale Limit Theory and its Applications.* Academic, New York.

Hu, Z., Shao, Q.M., Wang, Q. (2009). Cramér type moderate deviations for the maximum of self-normalized sums. *Electron. J. Probab.*, **14**(41): 1181–1197.

Liu, W., Shao, Q.M., Wang, Q. (2013). Self-normalized Cramér type moderate deviations for the maximum of sums. *Bernoulli*, **19**(3): 1006–1027.

Jing, B.Y., Shao, Q.M., Wang, Q. (2003). Self-normalized Cramér-type large deviations for independent random variables. *Ann. Probab.*, **31**(4): 2167–2215.

Jing, B.Y., Liang, H.Y., Zhou, W. (2012). Self-normalized moderate deviations for independnet random variables. *Sci. China Math.*, **55**(11): 2297–2315.

Petrov, V.V. (1975). *Sums of Independent Random Variables.* Springer-Verlag. Berlin.

Shao, Q.M. (1997). Self-normalized large deviations. *Ann. Probab.*, **25**(1): 285–328.

Shao, Q.M. (1999). A Cramér type large deviation result for Student's $t-$statistic. *J. Theor. Probab.*, **12**(2): 385–398.

Shao, Q.M., Wang, Q.Y. (2013). Self-normalized limit theorems: A survey. *Probab. Surv.*, **10**: 69–93.

Shao, Q.M., Zhou, W.X. (2016). Cramér type moderate deviation theorems for self-normalized processes. *Bernoulli*, **22**(4): 2029–2079.

Wang, Q.Y., Jing, B.Y. (1999). An exponential nonuniform Berry-Esseen bound for self-normalized sums. *Ann. Probab.*, **27**(4): 2068–2088.

von Bahr, B., Esseen, C.G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.*, **36**(1): 299–303.

arXiv: arXiv:0000.0000

Supplement to "Self-normalized Cramér type moderate deviations for martingales"

Xiequan Fan[1,*] **Ion Grama**[2,**] **Quansheng Liu**[2,†] **and Qi-Man Shao**[3,‡]

[1] *Center for Applied Mathematics, Tianjin University, Tianjin 300072, China. e-mail:* *fanxiequan@hotmail.com

[2] *Université de Bretagne-Sud, LMBA, UMR CNRS 6205, Campus de Tohannic, 56017 Vannes, France. e-mail:* **ion.grama@univ-ubs.fr; †quansheng.liu@univ-ubs.fr

[3] *Department of Statistics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong. e-mail:* ‡qmshao@sta.cuhk.edu.hk

Abstract: We give detailed proofs for Propositions 3.1 and 3.2 in the article "Self-normalized Cramér type moderate deviations for martingales".

1. Proof of Proposition 3.1

Recall that we have the notation

$$\zeta_i(\lambda) = \lambda\xi_i - \lambda^2\xi_i^2/2, \ \ \eta_i(\lambda) = \zeta_i(\lambda) - \mathbf{E}_\lambda[\zeta_i(\lambda)|\mathcal{F}_{i-1}], \ \ Y_n(\lambda) = \sum_{i=1}^n \eta_i(\lambda),$$

and $Y(\lambda) = (Y_k(\lambda), \mathcal{F}_k)_{k=0,\ldots,n}$. For simplicity, we write ζ_i, η_i, Y_n, Y for $\zeta_i(\lambda)$, $\eta_i(\lambda)$, $Y_n(\lambda)$, $Y(\lambda)$, respectively. In the sequel, ϑ (different from θ) stands for real numbers satisfying $0 \le \vartheta \le 1$ and φ stands for the density function of the standard normal distribution.

Let $\Delta\langle Y\rangle_k = \mathbf{E}_\lambda[\eta_i^2|\mathcal{F}_{k-1}]$ and $\langle Y\rangle_k = \sum_{i\le k}\Delta\langle Y\rangle_i$. Notice that for $\rho \ge 0$,

$$|\eta_i|^{2+\rho} \le 2^{1+\rho}(|\zeta_i|^{2+\rho} + \mathbf{E}_\lambda[|\zeta_i||\mathcal{F}_{i-1}]^{2+\rho}).$$

Using (A4), Lemmas 3.2 and 3.5 of [Fan et al., 2017], we have for all $0 \le \lambda = o(\gamma_n^{-1})$,

$$\begin{aligned}
\mathbf{E}_\lambda[|\eta_i|^{2+\rho}|\mathcal{F}_{i-1}] &\le 2^{1+\rho}\mathbf{E}_\lambda[|\zeta_i|^{2+\rho} + \mathbf{E}_\lambda[|\zeta_i||\mathcal{F}_{i-1}]^{2+\rho}|\mathcal{F}_{i-1}] \\
&\le 2^{2+\rho}\mathbf{E}_\lambda[|\zeta_i|^{2+\rho}|\mathcal{F}_{i-1}] \\
&= 2^{2+\rho}\frac{\mathbf{E}[|\lambda\xi_i - \lambda^2\xi_i^2/2|^{2+\rho}\exp\{\lambda\xi_i - \lambda^2\xi_i^2/2\}|\mathcal{F}_{i-1}]}{\mathbf{E}[\exp\{\lambda\xi_i - \lambda^2\xi_i^2/2\}|\mathcal{F}_{i-1}]} \\
&\le c_\rho'\frac{\mathbf{E}[|\lambda\xi_i|^{2+\rho}|\mathcal{F}_{i-1}]}{1 + O(1)\lambda^{2+\rho}\gamma_n^\rho\,\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]} \\
&\le c_\rho'\frac{\lambda^{2+\rho}\gamma_n^\rho\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]}{1 + O(1)\lambda^{2+\rho}\gamma_n^\rho\,\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]} \\
&\le c_\rho\lambda^{2+\rho}\gamma_n^\rho\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]. \qquad (1.1)
\end{aligned}$$

ver. 2014/10/16 file: Supplemnt_to_Cramer_type_large_deviations_for_martingales-180206.tex date: August 13, 2018

Using Lemmas 3.2 and 3.5 of [Fan et al., 2017] again, we obtain for all $0 \leq \lambda = o(\gamma_n^{-1})$,

$$
\begin{aligned}
&\big|\Delta \langle Y \rangle_k - \lambda^2 \mathbf{E}[\xi_k^2|\mathcal{F}_{k-1}]\big| \\
&\leq \left| \frac{\mathbf{E}[\zeta_k^2 e^{\zeta_k}|\mathcal{F}_{k-1}]}{\mathbf{E}[e^{\zeta_k}|\mathcal{F}_{k-1}]} - \lambda^2 \mathbf{E}[\xi_k^2|\mathcal{F}_{k-1}] \right| + \left| \frac{\mathbf{E}[\zeta_k e^{\zeta_k}|\mathcal{F}_{k-1}]^2}{\mathbf{E}[e^{\zeta_k}|\mathcal{F}_{k-1}]^2} \right| \\
&= \frac{\big|\mathbf{E}[\zeta_k^2 e^{\zeta_k}|\mathcal{F}_{k-1}] - \lambda^2 \mathbf{E}[\xi_k^2|\mathcal{F}_{k-1}]\mathbf{E}[e^{\zeta_k}|\mathcal{F}_{k-1}]\big|}{|\mathbf{E}[e^{\zeta_k}|\mathcal{F}_{k-1}]|} + \left| \frac{\mathbf{E}[\zeta_k e^{\zeta_k}|\mathcal{F}_{k-1}]^2}{\mathbf{E}[e^{\zeta_k}|\mathcal{F}_{k-1}]^2} \right| \\
&\leq c_1 \Big(\mathbf{E}[|\lambda \xi_k|^{2+\rho}|\mathcal{F}_{k-1}] + (\lambda^2 \mathbf{E}[\xi_k^2|\mathcal{F}_{k-1}])^2 \Big) \\
&\leq c_2\, \lambda^{2+\rho} \gamma_n^\rho \mathbf{E}[\xi_k^2|\mathcal{F}_{k-1}]. \qquad (1.2)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
|\langle Y \rangle_n - \lambda^2| &\leq |\langle Y \rangle_n - \lambda^2 \langle M \rangle_n| + \lambda^2 |\langle M \rangle_n - 1| \\
&\leq c\, \lambda^{2+\rho} \gamma_n^\rho \langle M \rangle_n + \lambda^2 \delta_n^2. \qquad (1.3)
\end{aligned}
$$

Inequalities (1.1) and (1.3) show that the martingale Y satisfies the following conditions. For all $0 < \lambda = o(\gamma_n^{-1})$,

(B1) $\mathbf{E}_\lambda\big[|\eta_i/\lambda|^{2+\rho}\,\big|\mathcal{F}_{i-1}\big] \leq c_\rho\, \gamma_n^\rho\, \mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}]$;
(B2) $\big|\langle Y \rangle_n / \lambda^2 - 1\big| \leq c\, \lambda^\rho \gamma_n^\rho + \delta_n^2$.

For simplicity of notation, set $T = 1 + \delta_n^2$. We introduce a modification of the conditional variance of the martingale M as follows:

$$V_k = \langle M \rangle_k \mathbf{1}_{\{k<n\}} + T \mathbf{1}_{\{k=n\}}. \qquad (1.4)$$

It is easy to see that $V_0 = 0$ and $V_n = T$, and that $(V_k, \mathcal{F}_k)_{k=0,\dots,n}$ is a predictable process. Denote

$$\widetilde{\gamma}_n = \lambda^\rho \gamma_n^\rho + \gamma_n + \delta_n.$$

Let c_* be a constant depending only on ρ, whose exact value will be chosen later. Then

$$A_k = c_*^2 \widetilde{\gamma}_n^2 + T - V_k, \quad k = 1, ..., n,$$

is a non-increasing predictable process. For any fixed $u, x \in \mathbf{R}$, and $y > 0$, set, for brevity,

$$\Phi_u(x, y) = \Phi\Big(\frac{u - x}{\sqrt{y}}\Big). \qquad (1.5)$$

In the proof we make use of the following two lemmas from [Bolthausen, 1982].

Lemma 1.1. *Let X and Y be random variables. Then*

$$\sup_u \Big|\mathbf{P}\left(X \leq u\right) - \Phi\left(u\right)\Big| \leq c_1 \sup_u \Big|\mathbf{P}\left(X + Y \leq u\right) - \Phi\left(u\right)\Big| + c_2 \Big\|\mathbf{E}\left[Y^2|X\right]\Big\|_\infty^{1/2}.$$

Lemma 1.2. *Let $G(x)$ be an integrable function on $\mathbf{R}$ of bounded variation $||G||_V$, X be a random variable, and a, $b \neq 0$ be real numbers. Then*

$$\mathbf{E}\Big[G\left(\frac{X+a}{b}\right)\Big] \leq ||G||_V \sup_u \Big|\mathbf{P}\left(X \leq u\right) - \Phi\left(u\right)\Big| + ||G||_1 \Big|b\Big|,$$

where $||G||_1$ is the $L_1(\mathbf{R})$ norm of $G(x)$.

Let $\mathcal{N} = \mathcal{N}(0,1)$ be a standard normal random variable independent of Y_n. Using Lemma 1.1, we deduce that

$$\begin{aligned}
\sup_u \Big|\mathbf{P}_\lambda(Y_n/\lambda \leq u) - \Phi(u)\Big| &\leq c_0 \sup_u \Big|\mathbf{P}_\lambda(c_* \widetilde{\gamma}_n \mathcal{N} + Y_n/\lambda \leq u) - \Phi(u)\Big| + c_1 c_* \widetilde{\gamma}_n \\
&= c_1 \sup_u \Big|\mathbf{E}_\lambda[\Phi_u(Y_n/\lambda, A_n)] - \mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)]\Big| + c_2 \widetilde{\gamma}_n \\
&\leq c_1 \sup_u \Big|\mathbf{E}_\lambda[\Phi_u(Y_n/\lambda, A_n)] - \mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)]\Big| \\
&\quad + c_1 \sup_u \Big|\mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)] - \Phi(u)\Big| + c_2 \widetilde{\gamma}_n \\
&= c_1 \sup_u \Big|\mathbf{E}_\lambda[\Phi_u(Y_n/\lambda, A_n)] - \mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)]\Big| \\
&\quad + c_1 \sup_u \Big|\Phi\Big(\frac{u}{\sqrt{c_*^2 \widetilde{\gamma}_n^2 + T}}\Big) - \Phi(u)\Big| + c_2 \widetilde{\gamma}_n, \qquad (1.6)
\end{aligned}$$

where the last line follows from the fact that $Y_0 = 0$ and $A_0 = c_*^2 \widetilde{\gamma}_n^2 + T$. Because $T = 1 + \delta_n^2$, it is obvious that

$$\sup_u \Big|\Phi\Big(\frac{u}{\sqrt{c_*^2 \widetilde{\gamma}_n^2 + T}}\Big) - \Phi(u)\Big| \leq c_3\, \widetilde{\gamma}_n.$$

Thus, from (1.6),

$$\sup_u \Big|\mathbf{P}_\lambda(Y_n/\lambda \leq u) - \Phi(u)\Big| \leq c_1 \sup_u \Big|\mathbf{E}_\lambda[\Phi_u(Y_n/\lambda, A_n)] - \mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)]\Big| + c_4 \widetilde{\gamma}_n. \qquad (1.7)$$

For the first item on the right hand side of the last inequality, we have the following telescoping

$$\mathbf{E}_\lambda[\Phi_u(Y_n/\lambda, A_n)] - \mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)] = \mathbf{E}_\lambda\Big[\sum_{k=1}^n \Big(\Phi_u(Y_k/\lambda, A_k) - \Phi_u(Y_{k-1}/\lambda, A_{k-1})\Big)\Big].$$

Taking into account that $(\eta_i, \mathcal{F}_i)_{i=0,\dots,n}$ is a $\mathbf{P}_\lambda$-martingale and that

$$\frac{\partial^2}{\partial x^2}\Phi_u(x,y) = 2\frac{\partial}{\partial y}\Phi_u(x,y),$$

we deduce that

$$\mathbf{E}_\lambda[\Phi_u(Y_n/\lambda, A_n)] - \mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)] = I_1 + I_2 - I_3, \qquad (1.8)$$

where

$$
\begin{aligned}
I_1 &= \mathbf{E}_\lambda\Big[\sum_{k=1}^n\Big(\Phi_u(Y_k/\lambda, A_k) - \Phi_u(Y_{k-1}/\lambda, A_k) \\
&\qquad - \frac{\partial}{\partial x}\Phi_u(Y_{k-1}/\lambda, A_k)\frac{\eta_k}{\lambda} - \frac{1}{2}\frac{\partial^2}{\partial x^2}\Phi_u(Y_{k-1}/\lambda, A_k)\frac{\eta_k^2}{\lambda^2}\Big)\Big], \\
I_2 &= \frac{1}{2}\mathbf{E}_\lambda\Big[\sum_{k=1}^n \frac{\partial^2}{\partial x^2}\Phi_u(Y_{k-1}/\lambda, A_k)\Big(\Delta\langle Y\rangle_k/\lambda^2 - \Delta V_k\Big)\Big], \\
I_3 &= \mathbf{E}_\lambda\Big[\sum_{k=1}^n\Big(\Phi_u(Y_{k-1}/\lambda, A_{k-1}) - \Phi_u(Y_{k-1}/\lambda, A_k) - \frac{\partial}{\partial y}\Phi_u(Y_{k-1}/\lambda, A_k)\Delta V_k\Big)\Big].
\end{aligned}
$$

Next, we estimate I_1, I_2, and I_3. To shorten the notations, denote

$$T_{k-1} = \frac{1}{\sqrt{A_k}}\Big(u - \frac{Y_{k-1}}{\lambda}\Big).$$

a) *Control of* I_1. Assume that f is a three times differentiable function on $\mathbf{R}$. By Taylor's expansion, it is easy to see that for any $|\Delta x| \leq 1$,

$$
\begin{aligned}
\Big|f(x+\Delta x) - f(x) - f'(x)\Delta x - \frac{1}{2}f''(x)(\Delta x)^2\Big| &= \Big|\frac{1}{6}f'''(x+\vartheta\Delta x)(\Delta x)^3\Big| \\
&\leq |f'''(x+\vartheta\Delta x)|\,|\Delta x|^{2+\rho},
\end{aligned}
$$

and for any $|\Delta x| > 1$,

$$
\begin{aligned}
&\Big|f(x+\Delta x) - f(x) - f'(x)\Delta x - \frac{1}{2}f''(x)(\Delta x)^2\Big| \\
&= \Big|\frac{1}{2}f''(x+\vartheta_1\Delta x)(\Delta x)^2 - \frac{1}{2}f''(x)(\Delta x)^2\Big| \\
&\leq \frac{1}{2}\Big(|f''(x+\vartheta_1\Delta x)| + |f''(x)|\Big)\,|\Delta x|^2 \\
&\leq |f''(x+\vartheta\Delta x)|\,|\Delta x|^2 \\
&\leq |f''(x+\vartheta\Delta x)|\,|\Delta x|^{2+\rho}.
\end{aligned}
$$

Taking $f(x) = \Phi(x)$, $x = T_{k-1}$ and $\Delta x = \frac{\eta_k}{\lambda\sqrt{A_k}}$, we have

$$
\begin{aligned}
|I_1| &\leq \mathbf{E}_\lambda\Big[\sum_{k=1}^n F\Big(T_{k-1} + \frac{\vartheta_1\eta_k}{\lambda\sqrt{A_k}}\Big)\Big|\frac{\eta_k}{\lambda\sqrt{A_k}}\Big|^{2+\rho}\mathbf{1}_{\{|\eta_k/\lambda\sqrt{A_k}|\leq 1+|T_{k-1}|/2\}}\Big] \\
&\quad + \mathbf{E}_\lambda\Big[\sum_{k=1}^n\Big|\Phi_u(Y_k/\lambda, A_k) - \Phi_u(Y_{k-1}/\lambda, A_k) - \frac{\partial}{\partial x}\Phi_u(Y_{k-1}/\lambda, A_k)\frac{\eta_k}{\lambda} \\
&\qquad - \frac{1}{2}\frac{\partial^2}{\partial x^2}\Phi_u(Y_{k-1}/\lambda, A_k)\frac{\eta_k^2}{\lambda^2}\Big|\mathbf{1}_{\{|\eta_k/\lambda\sqrt{A_k}|>1+|T_{k-1}|/2\}}\Big] \\
&:= I_{11} + I_{12}, \qquad (1.9)
\end{aligned}
$$

where

$$F(t) = \max\Big\{ \Big|\Phi'''(t)\Big|, \Big|\Phi''(t)\Big| \Big\}.$$

To bound the right hand side of (1.9), we distinguish two cases as follows.

Case 1: $|\eta_k/\lambda\sqrt{A_k}| \leq 1 + |T_{k-1}|/2$. By the inequality $F(t) \leq \varphi(t)(1+t^2)$, it follows that

$$\begin{aligned} F\Big(T_{k-1} + \frac{\vartheta_1 \eta_k}{\lambda\sqrt{A_k}}\Big) &\leq \varphi\Big(T_{k-1} + \frac{\vartheta_1 \eta_k}{\lambda\sqrt{A_k}}\Big)\bigg(1 + \Big(T_{k-1} + \frac{\vartheta_1 \eta_k}{\lambda\sqrt{A_k}}\Big)^2\bigg) \\ &\leq g_1(T_{k-1}), \end{aligned}$$

where

$$g_1(z) = \sup_{|t-z|\leq 1+|z|/2} \varphi(t)(1+t^2).$$

It is easy to see that $g_1(z)$ is non-increasing in $z \geq 0$. Because $g_1(z)$ is nonnegative,

$$I_{11} \leq \mathbf{E}_\lambda \bigg[\sum_{k=1}^n g_1(T_{k-1}) \Big| \frac{\eta_k}{\lambda\sqrt{A_k}} \Big|^{2+\rho} \mathbf{1}_{\{|\eta_k/\lambda\sqrt{A_k}|\leq 1+|T_{k-1}|/2\}} \bigg]. \tag{1.10}$$

Case 2: $|\eta_k/\lambda\sqrt{A_k}| > 1 + |T_{k-1}|/2$. It is easy to see that

$$\begin{aligned} &\Big|\Phi(x+\Delta x) - \Phi(x) - \Phi'(x)\Delta x - \frac{1}{2}\Phi''(x)(\Delta x)^2\Big| \\ &= \Big(\Big| \frac{\Phi(x+\Delta x) - \Phi(x)}{|\Delta x|^{2+\rho}} \Big| + |\Phi'(x)| + \Big|\frac{1}{2}\Phi''(x)\Big| \Big) |\Delta x|^{2+\rho} \\ &\leq \Big(4\Big| \frac{\Phi(x+\Delta x) - \Phi(x)}{(2+|x|)^2} \Big| + |\Phi'(x)| + |\Phi''(x)| \Big) |\Delta x|^{2+\rho} \\ &\leq \Big(\frac{c_1}{(2+|x|)^2} + |\Phi'(x)| + |\Phi''(x)| \Big) |\Delta x|^{2+\rho} \\ &\leq \frac{c_2}{(2+|x|)^2} |\Delta x|^{2+\rho} \end{aligned}$$

for $|\Delta x| > 1 + |x|/2$. Because $|\Phi''(t)| \leq 2$, it follows that

$$I_{12} \leq \mathbf{E}_\lambda \bigg[\sum_{k=1}^n g_2(T_{k-1}) \Big| \frac{\eta_k}{\lambda\sqrt{A_k}} \Big|^{2+\rho} \mathbf{1}_{\{|\eta_k/\lambda\sqrt{A_k}|> 1+|T_{k-1}|/2\}} \bigg], \tag{1.11}$$

where

$$g_2(z) = \frac{c_2}{(2+|z|)^2}.$$

Set

$$G(t) = g_1(t) + g_2(t).$$

It follows that

$$|I_1| \leq I_{11} + I_{12} \leq \mathbf{E}_\lambda \bigg[\sum_{k=1}^n G(T_{k-1}) \Big| \frac{\eta_k}{\lambda\sqrt{A_k}} \Big|^{2+\rho} \bigg]. \tag{1.12}$$

Now we consider the conditional expectation of $|\eta_k|^{2+\rho}$. Using condition (B1), we have

$$\mathbf{E}_\lambda[|\eta_k|^{2+\rho}|\mathcal{F}_{k-1}] \leq c_\rho\, \lambda^{2+\rho}\, \gamma_n^\rho\, \Delta\langle M\rangle_k,$$

where $\Delta\langle M\rangle_k = \langle M\rangle_k - \langle M\rangle_{k-1}$. From the definition of the process V, it follows that $\Delta\left\langle M\right\rangle_k = \Delta V_k = V_k - V_{k-1}, 1 \leq k < n$, and $\Delta\left\langle M\right\rangle_n \leq \Delta V_n$, and that

$$\mathbf{E}_\lambda[|\eta_k|^{2+\rho}|\mathcal{F}_{k-1}] \leq c_\rho \lambda^{2+\rho} \gamma_n^\rho\, \Delta V_k. \tag{1.13}$$

Returning to (1.12), by inequality (1.13), we get

$$|I_1| \leq J_1, \tag{1.14}$$

where

$$J_1 = c_\rho\, \gamma_n^\rho \mathbf{E}_\lambda\Big[\sum_{k=1}^n \frac{1}{A_k^{1+\rho/2}} G\left(T_{k-1}\right) \Delta V_k\Big]. \tag{1.15}$$

We introduce the time change τ_t as follows. For any real $t \in [0, T]$,

$$\tau_t = \min\{k \leq n : V_k > t\}, \quad \text{where} \quad \min \emptyset = n. \tag{1.16}$$

Let $(\sigma_k)_{k=1,\dots,n+1}$ be the increasing sequence of moments when the increasing stepwise function τ_t, $t \in [0, T]$, has jumps. It is clear that $\Delta V_k = \int_{[\sigma_k,\sigma_{k+1})} dt$, and that $k = \tau_t$ for $t \in [\sigma_k, \sigma_{k+1})$. Because $\tau_T = n$, we have

$$\begin{aligned}\sum_{k=1}^n \frac{1}{A_k^{1+\rho/2}} G\left(T_{k-1}\right) \Delta V_k &= \sum_{k=1}^n \int_{[\sigma_k,\sigma_{k+1})} \frac{1}{A_{\tau_t}^{1+\rho/2}} G\left(T_{\tau_t -1}\right) dt \\ &= \int_0^T \frac{1}{A_{\tau_t}^{1+\rho/2}} G\left(T_{\tau_t -1}\right) dt.\end{aligned}$$

Set $a_t = c_*^2 \widetilde{\gamma}_n^2 + T - t$. Because $\Delta V_{\tau_t} \leq \gamma_n^2 + 2\delta_n^2$ (cf., Lemma 3.5 of [Fan et al., 2017] and (1.4)), we see that

$$t \leq V_{\tau_t} \leq V_{\tau_t -1} + \Delta V_{\tau_t} \leq t + \gamma_n^2 + 2\delta_n^2, \quad t \in [0, T]. \tag{1.17}$$

Assume that $c_* \geq 4$. We have

$$\frac{1}{2}\, a_t \leq c_*^2 \widetilde{\gamma}_n^2 + T - (t + \gamma_n^2 + 2\delta_n^2) \leq A_{\tau_t} = c_*^2 \widetilde{\gamma}_n^2 + T - V_{\tau_t} \leq a_t, \quad t \in [0, T]. \tag{1.18}$$

Notice that $G(z)$ is symmetric, and is non-increasing in $z \geq 0$. From (1.15), (1.18) implies that

$$J_1 \leq c_\rho\, \gamma_n^\rho \int_0^T \frac{1}{a_t^{1+\rho/2}} \mathbf{E}_\lambda\Big[G\Big(\frac{u - Y_{\tau_t -1}/\lambda}{a_t^{1/2}}\Big)\Big] dt. \tag{1.19}$$

By Lemma 1.2, it is easy to see that

$$\mathbf{E}_\lambda\left[G\left(\frac{u - Y_{\tau_t -1}/\lambda}{a_t{}^{1/2}}\right)\right] \leq c_{\rho,1} \sup_z \Big|\mathbf{P}_\lambda(Y_{\tau_t -1}/\lambda \leq z) - \Phi(z)\Big| + c_{\rho,2}\sqrt{a_t}. \tag{1.20}$$

Because $V_{\tau_t-1}=V_{\tau_t}-\Delta V_{\tau_t}$, $V_{\tau_t}\geq t$ and $\Delta V_{\tau_t}\leq \gamma_n^2+2\,\delta_n^2$, we get

$$V_n-V_{\tau_t-1}\leq V_n-V_{\tau_t}+\Delta V_{\tau_t}\leq T-t+c_*(\gamma_n^2+\delta_n^2)\leq a_t. \tag{1.21}$$

By (1.1), we have

$$\mathbf{E}_\lambda[\eta_k^2/\lambda^2|\mathcal{F}_{k-1}]\leq c_\rho\Delta\left\langle M\right\rangle_k.$$

Thus

$$\begin{aligned}
\mathbf{E}_\lambda\left[(Y_n/\lambda-Y_{\tau_t-1}/\lambda)^2|\mathcal{F}_{\tau_t-1}\right] &= \mathbf{E}_\lambda\Bigg[\sum_{k=\tau_t}^n \mathbf{E}_\lambda[\eta_k^2/\lambda^2|\mathcal{F}_{k-1}]\Bigg|\mathcal{F}_{\tau_t-1}\Bigg]\\
&\leq c_\rho\,\mathbf{E}_\lambda\Bigg[\sum_{k=\tau_t}^n \Delta\left\langle M\right\rangle_k\Bigg|\mathcal{F}_{\tau_t-1}\Bigg]\\
&= c_\rho\,\mathbf{E}_\lambda\big[\left\langle M\right\rangle_n-\left\langle M\right\rangle_{\tau_t-1}|\mathcal{F}_{\tau_t-1}\big]\\
&\leq c_\rho\,\mathbf{E}_\lambda[V_n-V_{\tau_t-1}|\mathcal{F}_{\tau_t-1}]\\
&\leq c_\rho\,a_t.
\end{aligned}$$

Then, by Lemma 1.1, we deduce that for any $t\in[0,T]$,

$$\sup_z\Big|\mathbf{P}_\lambda(Y_{\tau_t-1}/\lambda\leq z)-\Phi(z)\Big|\leq c_{\rho,3}\sup_z\Big|\mathbf{P}_\lambda(Y_n/\lambda\leq z)-\Phi(z)\Big|+c_{\rho,4}\sqrt{a_t}. \tag{1.22}$$

From (1.19), using (1.20) and (1.22), we obtain

$$J_1\leq c_{\rho,5}\,\gamma_n^\rho\int_0^T\frac{dt}{a_t^{1+\rho/2}}\sup_z\Big|\mathbf{P}_\lambda(Y_n/\lambda\leq z)-\Phi(z)\Big|+c_{\rho,6}\,\gamma_n^\rho\int_0^T\frac{dt}{a_t^{(1+\rho)/2}}. \tag{1.23}$$

By (1.21) and some elementary computations, we see that

$$\int_0^T\frac{dt}{a_t^{1+\rho/2}}\leq\int_0^T\frac{dt}{(c_*^2\widetilde{\gamma}_n^2+T-t)^{1+\rho/2}}\leq\frac{c_\rho}{c_*^\rho\gamma_n^\rho}, \tag{1.24}$$

and

$$\int_0^T\frac{dt}{a_t^{(1+\rho)/2}}\leq\begin{cases}c_\rho, & \text{if } \rho\in(0,1),\\ c\,|\ln\gamma_n|, & \text{if } \rho=1.\end{cases}$$

Then

$$|I_1|\leq J_1\leq\frac{c_{\rho,7}}{c_*^\rho}\sup_z\Big|\mathbf{P}(Y_n/\lambda\leq z)-\Phi(z)\Big|+c_{\rho,8}\,\widehat{\gamma}_n, \tag{1.25}$$

where

$$\widehat{\gamma}_n=\begin{cases}\lambda^\rho\gamma_n^\rho+\gamma_n^\rho+\delta_n, & \text{if } \rho\in(0,1),\\ \lambda\gamma_n+\gamma_n\,|\ln\gamma_n|+\delta_n, & \text{if } \rho=1.\end{cases}$$

b) *Control of* I_2. Set $\widetilde{G}(z)=\sup_{|v|\leq 2}\psi(z+v)$, where $\psi(z)=\varphi(z)(1+z^2)^{3/2}$. Because $\Delta A_k=-\Delta V_k$, we have $|I_2|\leq I_{2,1}+I_{2,2}$, where

$$\begin{aligned}
I_{2,1} &= \mathbf{E}_\lambda\Bigg[\sum_{k=1}^n\frac{1}{2A_k}\left|\varphi'\left(T_{k-1}\right)\left(\Delta V_k-\Delta\left\langle M\right\rangle_k\right)\right|\Bigg],\\
I_{2,2} &= \mathbf{E}_\lambda\Bigg[\sum_{k=1}^n\frac{1}{2A_k}\left|\varphi'\left(T_{k-1}\right)\left(\Delta\left\langle Y\right\rangle_k/\lambda^2-\Delta\left\langle M\right\rangle_k\right)\right|\Bigg].
\end{aligned}$$

We first deal with $I_{2,1}$. Because $|\varphi'(z)| \leq \widetilde{G}(z)$ for any real z, we have

$$|\varphi'(T_{k-1})| \leq \widetilde{G}(T_{k-1}). \tag{1.26}$$

Notice that $0 \leq \Delta V_k - \Delta \langle M \rangle_k \leq 2\delta_n^2 \mathbf{1}_{\{k=n\}}$, $A_n = c_*^2 \widetilde{\gamma}_n^2$, and $c_* \geq 4$. Then, using (1.26), we get the estimations

$$I_{2,1} \leq \frac{c_{2,\rho}\delta_n^2}{c_*^2\widetilde{\gamma}_n^2} \mathbf{E}_\lambda[\widetilde{G}(T_{n-1})] \leq \frac{c_{2,\rho}}{c_*^2} \mathbf{E}_\lambda[\widetilde{G}(T_{n-1})],$$

and, by (1.20) with $G = \widetilde{G}$ and (1.22) with $t = T$,

$$|I_{2,1}| \leq \frac{c_{3,\rho}}{c_*} \sup_z \Big| \mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z) \Big| + c_{4,\rho}\widetilde{\gamma}_n.$$

We next consider $I_{2,2}$. By (1.2), we easily obtain the bound

$$\big| \Delta \langle Y \rangle_k / \lambda^2 - \Delta \langle M \rangle_k \big| \leq c_3 \, \lambda^\rho \gamma_n^\rho \Delta \langle M \rangle_k \leq c_3 \, \lambda^\rho \gamma_n^\rho \Delta V_k \ .$$

With this bound, we get

$$|I_{2,2}| \leq c_3 \, \lambda^\rho \gamma_n^\rho \mathbf{E}_\lambda \Big[\sum_{k=1}^n \frac{1}{2A_k} |\varphi'(T_{k-1})| \, \Delta V_k \Big].$$

Because $|\varphi'(z)| \leq \widetilde{G}(z)$, the right-hand side can be bounded exactly in the same way as J_1 in (1.15), with A_k replacing $A_k^{1+\rho/2}$. Similar to the proof of (1.23), we get

$$|I_{2,2}| \leq c_{5,\rho} \, \lambda^\rho \gamma_n^\rho \int_0^T \frac{dt}{a_t} \sup_z \Big| \mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z) \Big| + c_{6,\rho} \, \lambda^\rho \gamma_n^\rho \int_0^T \frac{dt}{a_t^{1/2}}.$$

By some elementary computations, we see that

$$\int_0^T \frac{dt}{a_t^{1/2}} \leq \int_0^T \frac{dt}{\sqrt{T-t}} \leq c_2,$$

and, taking into account that $a_t \geq c_*^2 \widetilde{\gamma}_n^2$,

$$\int_0^T \frac{dt}{a_t} \leq |\ln c_*^2 \widetilde{\gamma}_n^2| \leq c_\rho |\ln \lambda \gamma_n|.$$

Then

$$|I_{2,2}| \leq \frac{c_{7,\rho}}{c_*} \sup_z \Big| \mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z) \Big| + c_{8,\rho} \, \lambda^\rho \gamma_n^\rho.$$

Combining the bounds $I_{2,1}$ and $I_{2,2}$, we get

$$|I_2| \leq \frac{c_{9,\rho}}{c_*} \sup_z \Big| \mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z) \Big| + c_{10,\rho}\widetilde{\gamma}_n. \tag{1.27}$$

c) *Control of* I_3. By Taylor's expansion, it follows that

$$I_3 = \frac{1}{8}\mathbf{E}_\lambda\bigg[\sum_{k=1}^n \frac{1}{(A_k - \vartheta_k \Delta A_k)^2}\varphi'''\left(\frac{u - Y_{k-1}/\lambda}{\sqrt{A_k - \vartheta_k\Delta A_k}}\right)\Delta A_k^2\bigg].$$

Because $|\Delta A_k| = \Delta V_k \leq \gamma_n^2 + 2\,\delta_n^2$ and $c_* \geq 4$, we have

$$A_k \leq A_k - \vartheta_k \Delta A_k \leq c_*^2\widetilde{\gamma}_n^2 + T - V_k + 2\gamma_n^2 \leq 2A_k. \tag{1.28}$$

Using (1.28) and the inequalities $|\varphi'''(z)| \leq \widetilde{G}(z)$, we obtain

$$|I_3| \leq c\,(\gamma_n^2 + 2\,\delta_n^2)\mathbf{E}_\lambda\bigg[\sum_{k=1}^n \frac{1}{A_k^2}\,\widetilde{G}\left(\frac{T_{k-1}}{\sqrt{2}}\right)\Delta V_k\bigg].$$

Proceeding in the same way as for estimating J_1 in (1.15), we get

$$|I_3| \leq \frac{c_{11,\rho}}{c_*}\sup_z\Big|\mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z)\Big| + c_{12,\rho}\widetilde{\gamma}_n. \tag{1.29}$$

From (1.8), using (1.25), (1.27), and (1.29), we have

$$\Big|\mathbf{E}_\lambda[\Phi_u(Y_n/\lambda, A_n)] - \mathbf{E}_\lambda[\Phi_u(Y_0/\lambda, A_0)]\Big| \leq \frac{c'_{1,\rho}}{c_*^\rho}\sup_z\Big|\mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z)\Big| +\, c'_{2,\rho}\widehat{\gamma}_n.$$

Implementing the last bound in (1.7), we obtain

$$\sup_z\Big|\mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z)\Big| \leq \frac{c'_{3,\rho}}{c_*^\rho}\sup_z\Big|\mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z)\Big| + c'_{4,\rho}\,\widehat{\gamma}_n,$$

from which, choosing $c_*^\rho = \max\{2c'_{3,\rho}, 4^\rho\}$, we get

$$\sup_z\Big|\mathbf{P}_\lambda(Y_n/\lambda \leq z) - \Phi(z)\Big| \leq 2\,c'_{4,\rho}\,\widehat{\gamma}_n, \tag{1.30}$$

which gives the desired inequalities.

2. Proof of Proposition 3.2

Assume conditions (A1), (A2), and (A3). By Lemma 3.2 of [Fan et al., 2017], and condition (A3), it follows that for all $0 \leq \lambda = o(\kappa_n^{-1})$,

$$\begin{aligned}\mathbf{E}[\exp\{\lambda\xi_i - \lambda^2\xi_i^2/2\}|\mathcal{F}_{i-1}] &= 1 + O(1)\lambda^2\mathbf{E}[\xi_i^2|\mathcal{F}_{i-1}] \\ &= 1 + o(1).\end{aligned}$$

By an argument similar to the proof of (1.1), we get for all $0 \leq \lambda = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$\begin{aligned}\sum_{i=1}^n \mathbf{E}_\lambda[|\eta_i|^{2+\rho}|\mathcal{F}_{i-1}] &\leq c_\rho\sum_{i=1}^n \mathbf{E}[|\lambda\xi_i|^{2+\rho}|\mathcal{F}_{i-1}] \\ &\leq c_\rho\lambda^{2+\rho}\varepsilon_n^\rho\,.\end{aligned} \tag{2.1}$$

Similarly, we have for all $0 \leq \lambda = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

$$|\langle Y \rangle_n - \lambda^2| \leq c\, \lambda^{2+\rho} \varepsilon_n^{\rho} + \lambda^2 \delta_n^2 \tag{2.2}$$

(see (1.3) for a similar argument). Thus, Y satisfies the following conditions. For all $0 \leq \lambda = o(\max\{\varepsilon_n^{-1}, \kappa_n^{-1}\})$,

(C1) $\sum_{i=1}^n \mathbf{E}_\lambda \big[|\eta_i/\lambda|^{2+\rho} \big| \mathcal{F}_{i-1} \big] \leq c_\rho\, \varepsilon_n^{\rho}$;
(C2) $\big| \langle Y \rangle_n / \lambda^2 - 1 \big| \leq c\, \lambda^{\rho} \varepsilon_n^{\rho} + \delta_n^2$.

In the proof we make use of the following lemma of [Joos, 1993].

Lemma 2.1. *Let $\rho > 0$. Then*

$$\sup_x \Big| \mathbf{P}(M_n \leq x) - \Phi(x) \Big| \leq C_p \bigg(\Big(\sum_{i=1}^n \mathbf{E}|\xi_i|^{2+\rho} \Big)^{1/(3+\rho)} + || \langle M \rangle_n - 1| \, |_\infty^{1/2} \bigg). \tag{2.3}$$

Applying the last lemma to the martingale Y/λ, we obtain the desired inequalities.

References

Bolthausen, E. (1982). Exact convergence rates in some martingale central limit theorems. *Ann. Probab.*, **10**(3), 672–688.

Fan, X., Grama, I., Liu, Q., Shao, Q.M. (2017). Self-normalized Cramér type moderate deviations for martingales.

Joos, K. (1993). Nonuniform convergence rates in the central limit theorem for martingales. *Studia Sci. Math. Hungar.*, **28**(1-2), 145–158.